

JAN 10 2004

10/510903

SEQ ID NO:1

Size: 410

DNA--BAP-1

```
1  gcccgttgtc  tgtgtgtggg  actgaggggc  cccgggggcg  gtgggggctc  ccgggtggggg
61  cagcgggtggg  gagggagggc  ctggacatgg  cgctgagggg  ccgccccgcg  ggaagatgaa
121 taagggctgg  ctggaagctgg  agagcgaccc  aggcctcttc  accctgctcg  tggaagattt
181 cggtgtcaag  ggggtgcaag  tggaggagat  ctacgacctt  cagagcaaat  gtcagggccc
241 tgtatatgga  tttatcttcc  tgttcaaata  gatcgaagag  cgccgggtccc  ggcgaaagggt
301 ctctaccttg  gtggatgata  cgtccgtgat  tgatgatgat  attgtgaata  acatgttctt
361 tgcccaccag  ctgataccca  actcttgtgc  aactcatgcc  ttgctgagcg  tgctcctgaa
421 ctgcagcagc  gtggacctgg  gacccaccct  gagtcgcatg  aaggacttca  ccaagggttt
481 cagccctgag  agcaaaggat  atgcgattgg  caatgccccg  gagttggcca  aggcccataa
541 tagccatgcc  aggcccgagc  cagccacact  ccctgagaag  cagaatggcc  ttagtgcagt
601 gcggaccatg  gaggcgttcc  actttgtcag  ccatgtgcct  atcacaggcc  ggctctttga
661 gctggatggg  ctgaaggtdt  accccattga  ccatggggcc  tggggggagg  acgaggagtg
721 gacagacaag  gcccggcggg  tcatcatgga  gcgtatcggc  ctgcgacctg  caggggagcc
781 ctaccacgac  atccgcttca  acctgatggc  agtgggtgcc  gaccgcagga  tcaagtatga
841 ggccaggctg  catgtgctga  aggtgaaccg  tcagacagta  ctagaggctc  tgcagcagct
901 gataagagta  acacagccag  agctgattca  gaccacaag  tctcaagagt  cacagctgcc
961 tgaggagtcc  aagtcagcca  gcaacaagtc  cccgctggtg  ctggaagcaa  acagggcccc
1021 tgcagcctct  gagggcaacc  acacagatgg  tgcagaggag  gcggtctggt  catgcgca
1081 agccccatcc  cacagccctc  ccaacaaacc  caagctagtg  gtgaagcctc  caggcagcag
1141 cctcaatggg  gttcacccca  accccactcc  catgtccag  cggttgccgg  cctttctaga
1201 caatcacaat  tatgccaaat  cccccatgca  ggagggaaga  gacctggcgg  gacgtgtggg
1261 ccgcagccga  gttccagtcc  gcccacccca  gcagtactca  gatgatgagg  atgcatatga
1321 ggatgacgag  gaggatgacg  tgcagaacac  caactctgcc  cttaggtata  aggggaaggg
1381 aacaggggaag  ccaggggcat  tgagcggttc  tgctgatggg  caactgtcag  tgctgcagcc
1441 caacaccatc  aacgtcttgg  ctgagaagct  caaagagtcc  cagaaggacc  tctcaattcc
1501 tctgtccatc  aagactagca  gcggggctgg  gagtccggct  gtggcagtgc  ccacacactc
1561 gcagccctca  cccaccccca  gcaatgagag  tacagacacg  gcctctgaga  tcggcagtgc
1621 tttcaactcg  ccactgcgct  cgcctatccg  ctcagccaac  ccgacgcggc  cctccagccc
1681 tgtcacctcc  cacatctcca  aggtgctttt  tggagaggat  gacagctctg  tgcgtgttga
1741 ctgcatacgc  tacaaccgtg  ctgtccgtga  tctgggtcct  gtcatcagca  caggcctgct
1801 gcacctggct  gaggatgggg  tgctgagtcc  cctggcgctg  acagaggggt  ggaaggggtc
1861 ctgcgctccc  atcagaccaa  tccaaggcag  ccaggggtcc  agcagccag  tggagaagga
1921 ggtcgtggaa  gccacggaca  gcagagagaa  gacggggatg  gtgaggcctg  gcgagccctt
1981 gagtggggag  aaatactcac  ccaaggagct  gctggcactg  ctgaagtgtg  tggaggctga
2041 gattgcaaac  tatgaggcgt  gcctcaagga  ggaggtagag  aagaggaaga  agttcaagat
2101 tgatgaccag  agaaggaccc  acaactacga  tgagttcatc  tgcaccttta  tctccatgct
2161 ggctcaggaa  ggcatgctgg  ccaacctagt  ggagcagaac  atctccgtgc  ggcggcgcca
2221 aggggtcagc  atcggccggc  tccacaagca  gcggaagcct  gaccggcgga  aacgctctcg
2281 cccctacaag  gccaaagcgc  agtgaggact  gctggccctg  actctgcagc  ccactcttgc
2341 cgtgtggccc  tcaccagggt  ccttccctgc  cccacttccc  cttttccag  tattactgaa
2401 tagtcccagc  tggagagtcc  aggccctggg  aatgggagga  accaggccac  attccttcca
2461 tcgtgccctg  aggctgaca  cggcagatca  gccccatagt  gctcaggagg  cagcatctgg
2521 agttggggca  cagcgaggta  ctgcagcttc  ctccacagcc  ggctgtggag  cagcaggacc
2581 tggcccttct  gcctgggcag  cagaatatat  attttaccta  tcagagacat  ctatttttct
2641 gggtcccaac  ccaacatgcc  acctgttga  cataagttcc  tacctgacta  tgctttctct
2701 cctaggagct  gtccgtgtgg  gccaggtccc  ttgtatcatg  ccacggtccc  aactacaggg
2761 tcctagctgg  gggcctgggt  gggccctggg  ctctggggcc  tgctgctcta  gccccagcca
2821 ccagcctgtc  cctgttgtaa  ggaagccagg  tcttctctct  tcatctctct  taggagagtg
2881 ccaaactcag  ggaccagca  ctgggctggg  ttgggagtag  ggtgtcccag  tgggggtggg
2941 gtgagcaggc  tgctgggatc  ccatggcctg  agcagagcat  gtgggaactg  ttcagtggcc
3001 tgtgaactgt  ctctcttgtt  ctagccaggc  tgttcaagac  tgctctccat  agcaagggtc
3061 tagggctctt  cgccttcagt  gttgtggccc  tagctatggg  cctaaattgg  gctctaggtc
3121 tctgtccctg  gcgcttgagg  ctcagaagag  cctctgtcca  gcccctcagt  attaccatgt
```

FIG. 1 (1/2)

10/510903_

3181 ctccctctca ggggtagcag agacaggggt gcttatagga agctggcacc actcagctct
3241 tcctgctact ccagtttcct cagcctctgc aaggcactca ggggtgggga cagcaggatc
3301 aagacaaccc gttggagccc ctgtgttcca gaggacctga tgccaagggg taatgggccc
3361 agcagtgcct ctggagccca ggccccaaca cagcccatg gcctctgcca gatggctttg
3421 aaaaagggtga tccaagcagg cccctttatc tgtacatagt gactgagtgg ggggtgctgg
3481 caagtgtggc agctgcctct gggctgagca cagcttgacc cctctagccc ctgtaaatac
3541 tggatcaatg aatgaataaa actctcctaa gaatctcctg agaaaaaaaa aaaaaaaaaa

SEQ ID NO:2

Size: 729

PRT--BAP-1

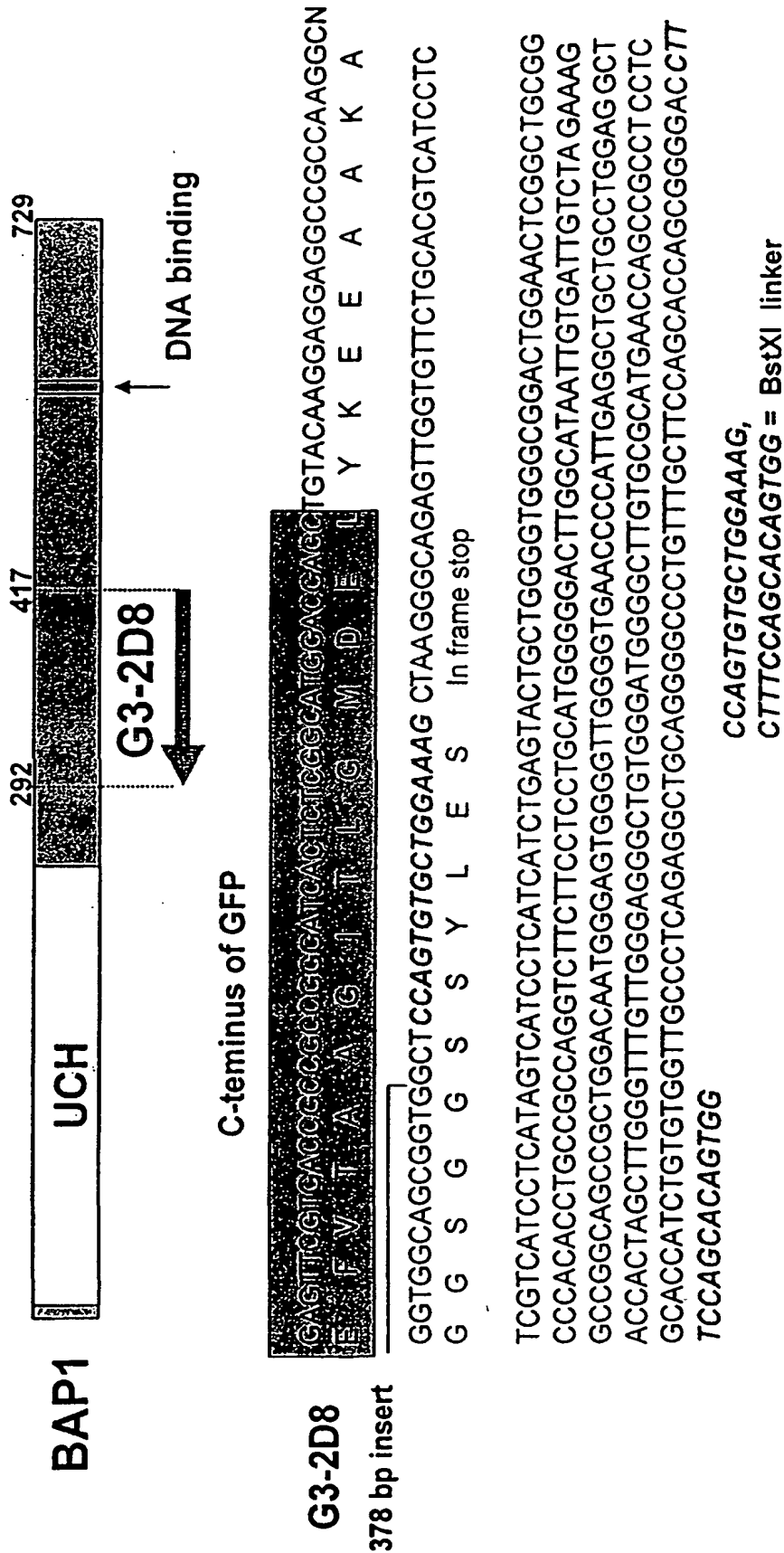
MNKGWLELESDPGLFTLLVEDFGVKGVQVEEIIDLQSKCQGPVYGFIFLFWIEERRSRKRVSTLVDDTSVIDDD
IVNNMFFAHQLIPNSCATHALLSVLLNCSSVDLGPTLSRMKDFTKGFSPESKGYAIGNAPELAKAHNSHARPEPR
HLPEKQNGLSAVRTMEAFHFVSYPITGRLFELDGLKVYPIDHGPWGEDEEWTDKARRVIMERIGLATAGEPYHD
IRFNLMAVVPDRRIKYEALHVLKVNROTVLEALQQLIRVTQPELIQTHKSQESQLPEESKSASNKSPVLEANR
APAASEGNHTDGABEAAGSCAQAPSHSPNPKPLVVKPPGSSLNGVHPNPTPIVQRLPAFLDNHNYAKSPMQEEE
DLAAGVGRSRVPVRPPQQYSDDDDYEDDEEDDVQNTNSALRYKGKGTGKPGALSGSADGQLSVLQPNNTINVLAE
KLKESQKDLSIPLSIKTSSGAGSPAVAVPTHSQPSPTPSNESTDTASEIGSAFNSPLRSPIRSANPTRPSSPPTS
HISKVLFGEDDSLRLVDCIRYNRAVRDLGPVISTGLLHLAEDGVLSPALTEGGKGSSPSIRPIQGSQGSSSPVE
KEVVEATDSREKTGMVRPGEPLSGEKYSPKELLALLKCVEAEIANYEACLKEEVEKRKKFKIDDQRRTHNYDEFI
CTFISMLAQEGMLANLVEQNISVRRRQGVSIGRLHKQRPDRRKRSRPHYKAKRQ

FIG. 1 (2/2)

10/510903

G3-2D8 / BRCA1-Associated Protein-1 (BAP1)

The G3-2D8 sequence is identical to BRCA1-Associated Protein-1 (BAP1), 729aa
Orientation: Antisense

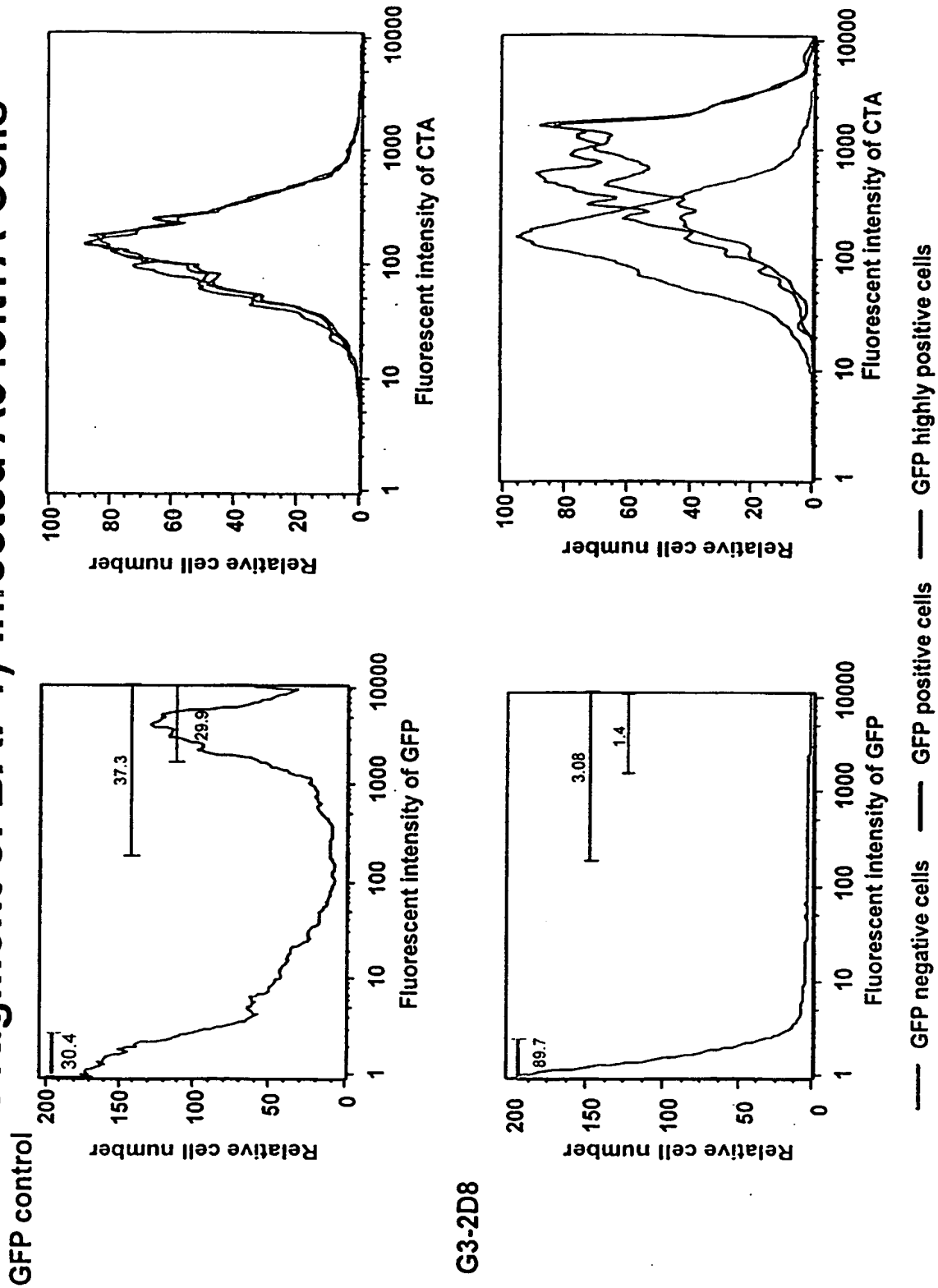


UCH(4-216): Ubiquitin carboxyl-terminal hydrolase, family 1,
DNA binding (625-640): 7kD DNA-binding domain

FIG. 2

10/510903

Cell Tracker Analysis of G3-2D8 (The Antisense Fragment of BAP1)-Infected A549.tTA Cells



10/510903

SEQ ID NO:3

Size: 437

DNA--NP95

```
1  CGACTCCTTA GAGCATGGCA TGGCTCAGAG GTGCTGGTAA AACTGATGGG GGTTTTTGCT
61 GTCCCTCCCC TCAGCGCCGA CACCATGTGG ATCCAGGTTT GGACCATGGA CGGGAGGCAG
121 ACCCACACGG TGGACTCGCT GTCCAGGCTG ACCAAGGTGG AGGAGCTGAG GCGGAAAGATC
181 CAGGAGCTGT TCCACGTGGA GCCAGGCCTG CAGAGGCTGT TCTACAGGGG CAAACAGATG
241 GAGGACGGCC ATACCCTCTT CGACTACGAG GTCCGCCTGA ATGACACCAT CCAGCTCCTG
301 GTCCGCCAGA GCCTCGTGCT CCCCCACAGC ACCAAGGAGC GGGACTCCGA GCTCTCCGAC
361 ACCGACTCCG GCTGCTGCCT GGGCCAGAGT GAGTCAGACA AGTCCTCCAC CCACGGCGAG
421 GCGGCCGCCG AGACTGACAG CAGGCCAGCC GATGAGGACA TGTGGGATGA GACGGAATTG
481 GGGCTGTACA AGGTCAATGA GTACGTCGAT GCTCGGGACA CGAACATGGG GGCGTGGTTT
541 GAGGCGCAGG TGGTCAGGGT GACGCGGAAG GCCCCTCCC GGGACGAGCC CTGCAGCTCC
601 ACGTCCAGGC CGGCGCTGGA GGAGGACGTC ATTTACCACG TGAAATACGA CGACTACCCG
661 GAGAACGGCG TGGTCCAGAT GAACTCCAGG GACGTCGAG CGCGCGCCCG CACCATCATC
721 AAGTGGCAGG ACCTGGAGGT GGGCCAGGTG GTCATGCTCA ACTACAACCC CGACAACCCC
781 AAGGAGCGGG GCTTCTGGTA CGACGCGGAG ATCTCCAGGA AGCGCGAGAC CAGGACGGCG
841 CGGGAACCTT ACGCCAACGT GGTGCTGGGG GATGATTCTC TGAACGACTG TCGGATCATC
901 TTCGTGGACG AAGTCTTCAA GATTGAGCGG CCGGGTGAAG GGAGCCCCAT GGTTGACAAC
961 CCCATGAGAC GGAAGAGCGG GCCGTCTGTC AAGCACTGCA AGGACGACGT GAACAGACTC
1021 TGCCGGGTCT GCGCCTGCCA CCTGTGCGGG GGCCGGCAGG ACCCCGACAA GCAGCTCATG
1081 TGCGATGAGT GCGACATGGC CTTCACATC TACTGCCTGG ACCCGCCCCT CAGCAGTGTT
1141 CCCAGCGAGG ACGAGTGTA CTGCCCTGAG TGCCGGAATG ATGCCAGCGA GGTGGTACTG
1201 GCGGGAGAGC GGCTGAGAGA GAGCAAGAAG AAGGCGAAGA TGGCCTCGGC CACATCTGCC
1261 TCACAGCGGG ACTGGGGCAA GGGCATGGCC TGTGTGGGCC GCACCAAGGA ATGTACCATC
1321 GTCCCGTCCA ACCACTACGG ACCCATCCCC GGGATCCCCG TGGGCACCAT GTGGCGGTTT
1381 CGAGTCCAGG TCAGCGAGTC GGGTGTCCAT CGGCCCCACG TGGCTGGCAT ACACGGCCGG
1441 AGCAACGACG GAGCGTACTC CCTAGTCCCT GCGGGGGGCT ATGAGGATGA CGTGGACCAT
1501 GGGAAATTTT TCACATACAC GGGTAGTGGT GGTGCGAGATC TTTCCGGCAA CAAGAGGACC
1561 GCGGAACAGT CTTGTGATCA GAAACTCACC AACACCAACA GGGCGCTGGC TCTCAACTGC
1621 TTTGCTCCCA TCAATGACCA AGAAGGGGCG GAGGCCAAGG ACTGGCGGTC GGGGAAGCCG
1681 GTCAGGGTGG TGCGCAATGT CAAGGGTGGC AAGAATAGCA AGTACGCCCC CGCTGAGGGC
1741 AACCCTACG ATGGCATCTA CAAGGTTGTG AAATACTGGC CCGAGAAGGG GAAGTCCGGG
1801 TTTCTCGTGT GCGCTACCT TCTGCGGAGG GACGATGATG AGCCTGGCCC TTGGACGAAG
1861 GAGGGGAAGG ACCGGATCAA GAAGCTGGGG CTGACCATGC AGTATCCAGA AGGCTACCTG
1921 GAAGCCCTGG CCAACCGAGA GCGAGAGAAG GAGAACAGCA AGAGGGAGGA GGAGGAGCAG
1981 CAGGAGGGGG GCTTCGCGTC CCCCAGGACG GGCAAGGGCA AGTGGAAGCG GAAGTCGGCA
2041 GGAGGTGGCC CGAGCAGGGC CGGGTCCCCG CGCCGGACAT CCAAGAAAAC CAAGGTGGAG
2101 CCCTACAGTC TCACGGCCCC GCAGAGCAGC CTCATCAGAG AGGACAAGAG CAACGCCAAG
2161 CTGTGGAATG AGGTCTGGC GTCACTCAAG GACCGGCCGG CGAGCGGCAG CCCGTTCCAG
2221 TTGTTCTGTA GTAAAGTGGA GGAGACGTTT CAGTGTATCT GCTGTGAGG GCTGGTGTTC
2281 CGGCCCATCA CGACCGTGTG CCAGCACAAC GTGTGCAAGG ACTGCCTGGA CAGATCCTTT
2341 CGGGCACAGG TGTTCACTG CCCTGCCTGC CGCTACGACC TGGGCCGAG CTATGCCATG
2401 CAGGTGAACC AGCCTCTGCA GACCGTCCCT AACCAGCTCT TCCCCGGCTA CGGCAATGGC
2461 CGGTGATCTC CAAGCACTTC TCGACAGGCG TTTTGCTGAA AACGTGTCGG AGGGCTCGTT
2521 CATCGGCACT GATTTTGTTC TTAGTGGGCT TAACTTAAAC AGGTAGTGTT TCCTCCGTTT
2581 CCTAAAAGG TTTGTCTTCC TTTTTTTTTT TTTTATTTT TCAAATCTAT ACATTTTCAG
2641 GAATTTATGT ATTCTGGCTA AAAGTTGGAC TTCTCAGTAT TGTGTTTAGT TCTTTGAAAA
2701 CATAAAAGCC TGCAATTTCT CGACAAAACA ACACAAGATT TTTTAAAGAT GGAATCAGAA
2761 ACTACGTGGT GTGGAGGCTG TTGATGTTTC TGGTGTCAAG TTCTCAGAAG TTCTGCCAC
2821 CAACTCTTTA AGAAGGCGAC AGGATCAGTC CTTCTCTAGG GTTCTGGCCC CCAAGGTCAG
2881 AGCAAGCATC TTCCTGACAG CATTTTGTCA TCTAAAGTCC AGTGACATGG TTCCCCGTGG
2941 TGGCCCGTGG CAGCCCGTGG CATGGCGTGG CTCAGCTGTC TGTGTAAGTT GTTGCAAGGA
3001 AAAGAGGAAA CATCTCGGGC CTAGTTCAAA CCTTGCCCTC AAAGCCATCC CCCACCAGAC
3061 TGCTTAGCGT CTGAGATCCG CGTGAAAAGT CCTCTGCCCA CGAGAGCAGG GAGTTGGGGC
3121 CACGCAGAAA TGGCCTCAAG GGGACTCTGC TCCACGTGGG GCCAGGCGTG TGA CTGACGC
```

FIG. 4 (1/2)

10/510903

3181 TGTCCGACGA AGGCGGCCAC GGACGGACGC CAGCACACGA AGTCACGTGC AAGTGCCTTT
3241 GATTCGTTCC TTCTTTCTAA AGACGACAGT CTTTGTGTGT AGCACTGAAT TATTGAAAAT
3301 GTCAACCAGA TTCTAGAAAC TGCGGTCATC CAGTTCTTCC TGACACCGGA TGGGTGCTTG
3361 GGAACCGTTT GAGCCTTATA GATCATTTAC ATTCAATTTT TTAACTCAG CAAGTGAGAA
3421 CTTACAAGAG GGTTTTTTTT TAATTTTTTT TTCTCTTAAT GAACACATTT TCTAAATGAA
3481 TTTTTTTTGT AGTTACTGTA TATGTACCAA GAAAGATATA ACGTTAGGGT TTGGTTGTTT
3541 TTGTTTTTGT ATTTTTTTTC TTTTGAAAGG GTTTGTTAAT TTTTCTAATT TTACCAAAGT
3601 TTGCAGCCTA TACCTCAATA AAACAGGGAT ATTTTAAATC ACATACCTGC AGACAAACTG
3661 GAGCAATGTT ATTTTTTAAAG GGTTTTTTTT ACCTCCTTAT TCTTAGATTA TTAATGTATT
3721 AGGGAAGAAT GAGACAATTT TGTGTAGGCT TTTTCTAAAG TCCAGTACTT TGTCCAGATT
3781 TTAGATTCTC AGAATAAATG TTTTTCACAG ATTGAAAAAA AAAAAAAA

SEQ ID NO:4

Size: 135

PRT--NP95

MWIVRTMDGRQTHTVDSLRLTKVEELRRKIQELFHVEPGLQRLFYRGKQMEDGHTLFDYEVRLNDTIQLLVRO
SLVLPSTKERDSELSDTDSGCCLGQSESDKSSTHGEAAAETDSRPADEDMWDETELGLYKVNEYVDARDTNMGA
WFEAQVVRVTRKAPSRDEPCSSSTRPALEEDVIYHVKYDDYPENGVVQMNSRDVRRARARTIIKWQDLEVGQVVML
NYPNPNPKERGFWDYDAEISRKRETRTARELYANVVLGDDSLNDCRIIFVDEVFKIERPGECSMPVDNPMRRKSGP
SCKHCKDDVNRLCRVCACHLCGGRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSSEDEWYCPECRNDASEVVLAGE
RLRESKKKAKMASATSSSQRDWKGKMACVGRTECTIVPSNHYGPIPGIPVGTMRFRVQVSESGVHRPHVAGIH
GRSNDGAYSLVLAGGYEDDVDHGNFFTYTGSGGRDLGNGKRTAEQSCDQKLTNTNRALALNCFAPINDQEGAEAK
DWRSGKPVVRNVKGGKNSKYAPAEGNRYDGIYKVVKYWPEKGKSGFLVWRYLLRRDDDEPGPWTKGKDKRIKK
LGLTMQYPEGYLEALANREREREKENSKEEEEQEGGFASPRTGKGKWKRSAGGGPSRAGSPRRTSKTKVEPYS
LTAQQSSLIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCCQELVFRPITTVQCQHNVCCKDCLDR
SFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR

FIG. 4 (2/2)

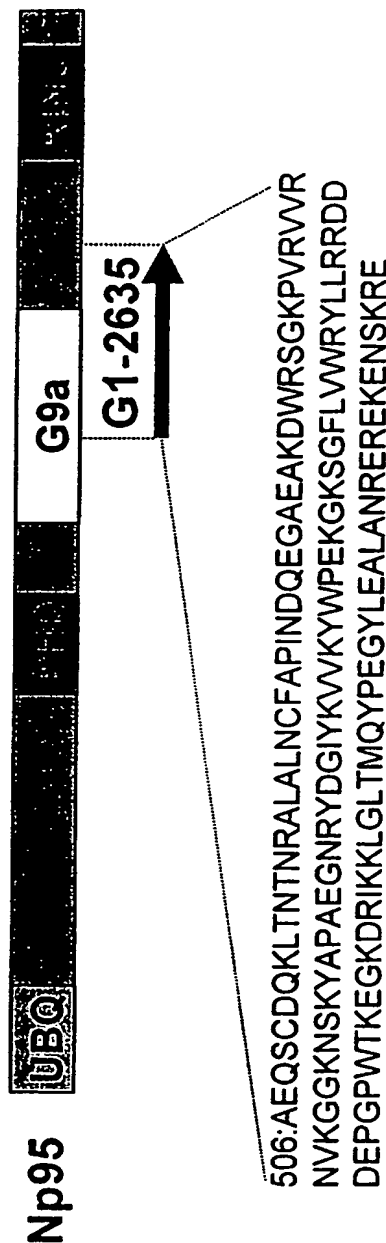
10/510903

G1-2635 / Np95

The G1-2635 sequence is identical to a nuclear zinc finger protein, Np95, 793aa

Orientation of cDNA: Sense

Pfam HMM search was done at the Washington University web site



UBQ(14-89): Ubiquitin like domain,
 PHD(330-379): PHD-Zn finger, It could be important for the assembly or activity of multicomponent complexes
 G9a(427-599): It is found in a nuclear protein associated with cell proliferation
 RING(737-775): Zinc finger, C3HC4 type (RING finger), E3 ubiquitin-protein ligase activity is intrinsic to the RING domain of c-Cbl and is likely to be a general function of this domain; Various RING fingers exhibit binding to E2 ubiquitin-conjugating enzymes

FIG. 5

Cell Tracker Analysis of G1-2635 (The Fragment of Np95)-Infected A549.tTA Cells

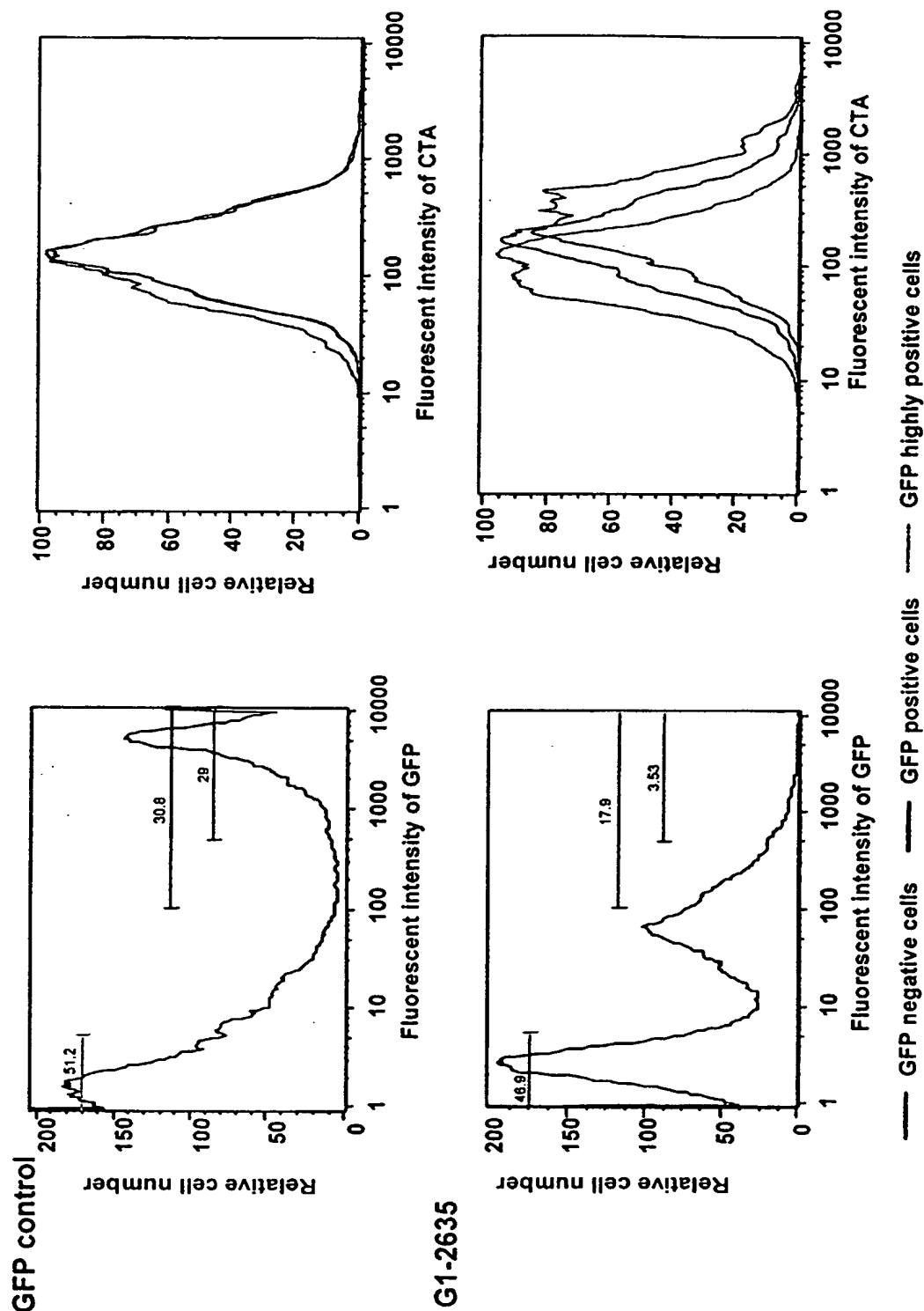


FIG. 6

FIG. 7 (1/5)

10/510903



Nucleotide

PubMed

Nucleotide

Protein

Genome

Structure

PopSet

Taxonomy

OMIM

8

Search **Nucleotide** for

1: NM_000135. Homo sapiens
Fanc...[gi:4503654]

Related Sequences, OMIM, Protein, PubMed, Taxonomy,
UniSTS, LinkOut

LOCUS NM_000135 5503 bp mRNA linear PRI 05-JUL-2001
DEFINITION Homo sapiens Fanconi anemia, complementation group A (FANCA), mRNA.
ACCESSION NM_000135
VERSION NM_000135.1 GI:4503654

FANCA

KEYWORDS

SOURCE

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

- 1 (bases 1 to 5503)
Pronk JC, Gibson RA, Savoia A, Wijker M, Morgan NV, Melchionda S, Ford D, Temtamy S, Ortega JJ, Jansen S and et al.
Localisation of the Fanconi anaemia complementation group A gene to chromosome 16q24.3
Nat. Genet. 11 (3), 338-340 (1995)
96042586
7581462
- 2 (bases 1 to 5503)
Lo Ten Foe, J.R., Rooimans, M.A., Bosnoyan-Collins, L., Alon, N., Wijker, M., Parker, L., Lightfoot, J., Carreau, M., Callen, D.F., Savoia, A., Cheng, N.C., Van Berkel, C.G.M., Strunk, M.H.P., Gille, J.J.P., Pals, G., Kruyt, F.A.E., Pronk, J.C., Arwert, F., Buchwald, M. and Joenje, H.
Expression cloning of a cDNA for the major Fanconi anaemia gene, FAA
Nat. Genet. 14 (3), 320-323 (1996)
97051928
- 3 (bases 1 to 5503)
Ianzano L, D'Apolito M, Centra M, Savino M, Levran O, Auerbach AD, Cleton-Jansen AM, Doggett NA, Pronk JC, Tipping AJ, Gibson RA, Mathew CG, Whitmore SA, Apostolou S, Callen DF, Zelante L and Savoia A.
The genomic organization of the Fanconi anemia group A (FAA) gene
Genomics 41 (3), 309-314 (1997)
97312685
9169126
- 4 (bases 1 to 5503)
Joenje H, Oostra AB, Wijker M, di Summa FM, van Berkel CG, Rooimans MA, Ebell W, van Weel M, Pronk JC, Buchwald M and Arwert F.
Evidence for at least eight Fanconi anemia genes
Am. J. Hum. Genet. 61 (4), 940-944 (1997)
98018453
9382107
- 5 (bases 1 to 5503)
Kupfer GM, Naf D, Suliman A, Pulsipher M and D'Andrea AD.
The Fanconi anaemia proteins, FAA and FAC, interact to form a

nuclear complex
JOURNAL Nat. Genet. 17 (4), 487-490 (1997)
MEDLINE 98061104
PUBMED 9398857
COMMENT PROVISIONAL REFSEQ: This record has not yet been subject to final
NCBI review. The reference sequence was derived from X99226.1.
FEATURES
 source
 Location/Qualifiers
 1..5503
 /organism="Homo sapiens"
 /isolate="healthy control"
 /db_xref="taxon:9606"
 /chromosome="16"
 /map="16q24.3"
 /clone="D"
 /cell_line="HSC93"
 /cell_type="lymphoblastoid"
 /clone_lib="pREP4"
 gene
 1..5503
 /gene="FANCA"
 /note="FA; FA1; FAA; FACA; FANCH"
 /db_xref="LocusID:2175"
 /db_xref="MIM:227650"
 /db_xref="MIM:603468"
 CDS
 32..4399
 /gene="FANCA"
 /function="acts with other genes to control FA pathway"
 /note="Fanconi anemia, complementation group H"
 /codon_start=1
 /db_xref="LocusID:2175"
 /db_xref="MIM:227650"
 /db_xref="MIM:603468"
 /product="Fanconi anemia, complementation group A"
 /protein_id="NP_000126.1"
 /db_xref="GI:4503655"
 /translation="MSDSWVPNSASGQDPGRRRAWAELLAGRVKREKYNPERAQKLK
ESAVRLLRSHQDLNALLLEVEGPLCKKLSLSKVIDCDSSEAYANHSSSFIGSALQDQA
SRLGVPVGILSAGMVASSVGQICTAPAETSHPVLLTVEQRKKLSSLLLEFAQYLLAHSM
FSRLSFCQELWKIQSSLLLEAVWHLHVQGIIVSLQELLESHPDMAVGSWLFRLNLCCLC
EQMEASCQHADVARAMLSDFVQMFVLRGFGQKNSDLRRTVEPEKMPQVTVQVLRMLIF
ALDALAAGVQEESSSTHKIVRCWFGVFSGHTLGSVISTDPLKRRFFSHTLTQILTHSPVL
KASDAVQMQREWSFARTHPLLTSLYRRLFVMLSABEELVGHLEQVLETOQEVHWQVLSF
VSALVVCFFPEAQQLLEDWVARLMAQAFESCQLDSMVTAFLLVVRQAALLEGPSAFLSYAD
WFKASFGSTRGYHGCSKKALVFLFTFLSELVFPFESPRYLQVHILHPPLVPKYSRLT
DYISLAKTRLADLKVSIENMGLEYEDLSSAGDITEPHSQALQDVEKAIMVFEHTGNIPV
TVMEASIFRRPYVSHFLPALLTPRVLPKVPDSRVAFIESLKRADKIPPSLYSTYCOA
CSAAEEKPEDAALGVRAEPNSABEPLGQLTAALGELRASMTDPSQRDVISQVAVISE
RLRAVLGHNEDDSSVEISKIQLSINTPRLEPREHIAVDLLLTSCQNLMAASSVAPPE
RQGPWAALFVRTMCGRVLPVAVLTRLCQLLRHQGPSLSAPHVLGLAALAVHLGESRSAL
PEVDVGPPAPGAGLPVPALFDSLLTCTRTDLSLFFCLKFCTAAISYSLCKFSSQSRDTL
CSCLSPLGIKKFQFLMFRLFSEARQPLSEEDVASLSWRPLHLPSADWQRAALSFWTHR
TFREVLKEEDVHLTYQDWLHLELEIQPEADALSDTERQDFHQWAIHEHFLPESSASGG
CDGDLQAACTILVNALMDFHQSSRSYDHSNSDLVFGGRTGNEDIISRLQEMVADLEL
QQDLIVPLGHTPSQEHLFEIFRRRLQALTSQWSVAASLQRORELLMYKRILLRLPSS
VLCGSSFQAEQPI TARCEQFFHLVNSEMRNFCSHGGALTQDITAHFRLGLLNACLRSR
DPSLMVDFILAKCQTKCPLILTSALVWWSLEPVLLCRWRRHCQSPLEQLKQEGR
QFASDFLSPEAASPAPNPDWLSAAALHFAIQQVREENIRKQKKLDCEREELLVFLFF
FSLMGLLSSHLTSNSTTDLPKAFHVCAAILCCLKRKISWLALFQLTESDLRLGRLLL
RVAPDQHTRLLPFAFYSLLSYFHEDAAIREEAFLHVAVDMYLLKLVQLFVAGDTSTVSP
PAGRSLELKGQGNPVELITKARLFLQLIPRCPKKFSFHVAAELLADRGDCDPEVSAAL
QSRQQAAPDADLSQEPHLF"

10/510903

FIG. 7 (3/5)

variation 48
/allele="A"
/allele="T"
/db_xref="dbSNP:1800282"

variation 1174
/allele="G"
/allele="T"
/db_xref="dbSNP:1800331"

variation 1321
/allele="A"
/allele="G"
/db_xref="dbSNP:1800332"

variation complement (1532)
/allele="C"
/allele="T"
/db_xref="dbSNP:2239359"

variation 3214
/allele="C"
/allele="T"
/db_xref="dbSNP:1800346"

variation 3685
/allele="A"
/allele="G"
/db_xref="dbSNP:1800358"

variation 4553
/allele="A"
/allele="G"
/db_xref="dbSNP:1230"

BASE COUNT 1208 a 1527 c 1492 g 1276 t
ORIGIN

```

1 agccgccgcc ggggctgtag gcgccaaggc catgtccgac tcgtgggtcc cgaactccgc
61 ctccggccag gacccagggg gccgccggag ggccctgggccc gagctgctgg cgggaaggggt
121 caagagggaa aaatataatc ctgaaagggc acagaaatta aaggaatcag ctgtgcgccct
181 cctgcgaagc catcaggacc tgaatgccct tttgcttgag gtagaagggtc cactgtgttaa
241 aaaattgtct ctacagcaaag tgattgactg tgacagttct gaggcctatg ctaatcattc
301 tagttcattt ataggctctg ctttgcagga tcaagcctca aggctggggg tccccgtggg
361 tattctctca gccgggatgg ttgcctctag cgtgggacag atctgcacgg ctccagcgga
421 gaccagtcac cctgtgctgc tgactgtgga gcagagaaag aagctgtctt ccctgttaga
481 gtttgcctcag tatttattgg cacacagtat gttctcccg ttttccttct gtcaagaatt
541 atggaaaata cagagttctt tgttgcttga agcgggtgtgg catcttcacg tacaaggcat
601 tgtgagcctg caagagctgc tggaaaagcca tcccagacatg catgctgtgg gatcgtggct
661 cttcaggaat ctgtgctgcc tttgtgaaca gatggaagca tcctgccagc atgctgacgt
721 cgccagggcc atgctttctg attttgttca aatgtttgtt ttgaggggat ttcagaaaaa
781 ctcagatctg agaagaactg tggagcttga aaaaatgccg caggtcacgg ttgatgtact
841 gcagagaatg ctgatttttg cacttgacgc tttggctgct ggagtacagg aggagtctc
901 cactcacaag atcgtgaggt gctggttcgg agtggtcagt ggacacacgc ttggcagtgt
961 aatttcaca gatcctctga agaggttctt cagtcatacc ctgactcaga tactcactca
1021 cagccctgtg ctgaaagcat ctgatgctgt tcagatgcag agagagtggg gctttgcgcg
1081 gacacaccct ctgctcacct cactgtaccg caggctcttt gtgatgctga gtgcagagga
1141 gttggttggc catttgcaag aagttctgga aacgcaggag gttcactggc agagagtgtc
1201 ctcccttggc tctgccctgg ttgtctgctt tccagaagcg cagcagctgc ttgaagactg
1261 ggtggcgcg tttgatggcc aggcattcga gagctgccag ctggacagca tggctactgc
1321 gttcctggtt gtgcgccagg cagcactgga gggcccctct gcgttcctgt catatgcaga
1381 ctggttcaag gcctcctttg ggagcacacg aggtaccat ggctgcagca agaaggccct
1441 ggtcttctctg tttacgttct tgtcagaact cgtgcctttt gagtctcccc ggtacctgca
1501 ggtgcacatt ctccaccac ccctggttcc cagcaagtac cgctccctcc tcacagacta
1561 catctcattg gccaaagacac ggctggccga cctcaagggt tctatagaaa acatgggact
1621 ctacgaggat ttgtcatcag ctggggacat tactgagccc cacagccaag ctcttcagga
1681 tgttgaaaag gccatcatgg tgtttgagca tacggggaac atcccagtc cgtcatgga
1741 ggccagcata ttcaggaggc cttactacgt gtcccacttc ctccccgcc tgctcacacc

```

10/510903

FIG. 7 (4/5)

```

1801 tgcagtgctc cccaaagtcc ctgactcccg tgtggcggtt atagagtctc tgaagagagc
1861 agataaaatc ccccatctc tgtactccac ctactgccag gcctgctctg ctgctgaaga
1921 gaagccagaa gatgcagccc tgggagtga ggcagaaccc aactctgctg aggagccct
1981 gggacagctc acagctgcac tgggagagct gagagcctcc atgacagacc ccagccagcg
2041 tgatgttata tcggcacagg tggcagtgat ttctgaaaga ctgagggctg tcttgggcca
2101 caatgaggat gacagcagcg ttgagatata aaagattcag ctgagcatca acacgccgag
2161 actggagcca cgggaacaca ttgctgtgga cctcctgctg acgtctttct gtcagaacct
2221 gatggctgcc tccagtgtcg ctcccccgga gaggcagggg ccctgggctg ccctcttctg
2281 gaggaccatg tgtggacgtg tgctccctgc agtgctcacc cggctctgcc agctgctcgg
2341 tcaccagggc ccgagcctga gtgcccaca tgtgctgggg ttggctgccc tggccgtgca
2401 cctgggtgag tccaggtctg cgctcccaga ggtggatgtg ggctcctctg cacttggtgc
2461 tggccttctc gtccctgcgc tctttgacag cctcctgacc tgtaggacga gggattcctt
2521 gttcttctgc ctgaaatttt gtacagcagc aatttcttac tctctctgca agtttcttct
2581 ccagtcacga gatactttgt gcagctgctt atctccaggc cttattaaaa agtttcagtt
2641 cctcatgttc agattgttct cagaggcccg acagcctctt tctgaggagg acgtagccag
2701 cctttcctgg agacccttgc accttccttc tgcagactgg cagagagctg ccctctctct
2761 ctggacacac agaaccttcc gagagggtgt gaaagaggaa gatgttcaact taacttacca
2821 agactggta cacctggagc tggaaattca acctgaagct gatgctcttt cagatactga
2881 acggcaggac ttccaccagt gggcgatcca tgagcacttt ctccctgagt cctcggttc
2941 agggggctgt tcaggagacc tgcaggctgc gtgtaccatt cttgtcaacg cactgatgga
3001 tttccaccaa agctcaagga gttatgacca ctgaaaaat tctgatttgg tcttgggtgg
3061 ccgcacagga aatgaggata ttatttccag attgcaggag atggtagctg acctggagct
3121 gcagcaagac ctcatagtgc ctctcgcca cacccttcc caggagcact tctcttttga
3181 gattttccgc agacggctcc aggtctgac aagcgggtgg agcgtggctg ccagccttca
3241 gagacagagg gagctgctaa tgtacaaacg gatcctcctc cgctgcctt cgtctgtcct
3301 ctgcggcagc agcttccagg cagaacagcc catcactgcc agatgcgagc agttcttcca
3361 ctgggtcaac tctgagatga gaaacttctg ctcccacgga ggtgccctga cacaggacat
3421 cactgccac ttcttcagg gctcctgtaa cgctgtctg cggagcagag acccctcctt
3481 gatggtcgac ttcatactgg ccaagtggca gcgaaatgc cccttaattt tgactgtgc
3541 tctggtgtgg tggccgagcc tggagcctgt gctgctctgc cgggtggagga gacactgcc
3601 gagcccgctg ccccggaac tgcagaagct acaagaaggc cggcagtttg ccagcgattt
3661 cctctccctt gaggtgcct cccagcacc caaccggac tggctctcag ctgctgcact
3721 gcactttgcg attcaacaag tcagggaaga aaacatcagg aagcagctaa agaagctgga
3781 ctgcgagaga gaggagctat tggttttcct tttcttcttc tcttgatgg gcctgctgtc
3841 gtcacatctg acctcaaata gcaccacaga cctgccaaag gctttccacg tttgtgcagc
3901 aatcctcgag tgtttagaga agaggaagat atcctggctg gcactcttct agttgacaga
3961 gagtgaactc aggtggggc ggctcctct cctgttggcc cggatcagc acacaggct
4021 gctgccttct gctttttaca gtcttctct ctacttccat gaagacggc ccatcaggga
4081 agaggccttc ctgcatgttg ctgtggacat gtacttgaag ctgggtccagc tcttcgtggc
4141 tggggataca agcacagttt cacctccagc tggcaggagc ctggagctca agggtcaggg
4201 caaccccgct gaactgataa caaaagctcg tctttttctg ctgcagttaa tacctcggtg
4261 cccgaaaaag agcttctcac acgtggcaga gctgctggct gatcgtgggg actgcgaccc
4321 agagggtgag gccgcctcc agagcagaca gcaggctgcc cctgacgctg acctgtccca
4381 ggagcctcat ctcttctgac gggacctgcc actgcacacc agcccagctc ccgtgtaa
4441 aatttattac aagcataaca tggagctctt gttgcactaa aaagtggatt acaaatctcc
4501 tcgactgctt tagtggggaa aggaatcaat tatttatgaa ctgtccggcc ccgagtcact
4561 cagcgttttg cggaaaaata accactggtc ccagagcaga ggaaggctac ttgagccgga
4621 caccaagccc gcctccagca ccaaggcggt gcagcaccct ccgacctcc catgcgggtg
4681 cacacgaagg gtgaggctga cacagccact gcggagtcca ggctgctaga ggtgctcatc
4741 ctactgccg tctcaggtg ggttcgggct tcaccgctg gccctctgtg gtcacagagg
4801 ggctcgggtg cccaggtggt ggttcgcct ccaggggcag ggccttctca tgggtctgtg
4861 tcagcgggtg caccatggac atgtgtacat tgaggttgtg ggccttctca aaccgcccgc
4921 cacactggtc acaggcaaag tccagctcag tctcagcctt gtgtttggtc atgtggtact
4981 tgagggatgc ccgtgcctg cactggaacc cacagacctc acacctgggg gacagaggca
5041 gataagaagg tgcagggcca cagcctggg aggggtcct gactcacact tactgcaaag
5101 gcttggctcc cgaatgtgc atttgggtga cgagaagggt cttccgctgc ttgaaggttt
5161 gtccacattc gtcacagata tagttccgca cctctgagag gggagagtcc agtgagttca
5221 ggccctgat gctccaacct cccgggggga cgacgatgac aatgtgaaac catcacagct
5281 gggaagacat ttctgcacat ggttcacat gcagtgggca caagcaagg gcctatgagg
5341 gcctcgttta ttaagatctt taaactgctt tatacactgt cactgggtt catcagctgt

```

10/510903

FIG. 7 (5/5)

5401 gtgcatttca ggatggtttt taaagaaacc tcagaaagct atttccttaa aaaaaaaaaa
5461 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaa

//

Revised: October 24, 2001.

Disclaimer | Write to the Help Desk
NCBI | NLM | NIH

FIG. 8 (1/4)

10/510903



Nucleotide

PubMed

Nucleotide

Protein

Genome

Structure

PopSet

Taxonomy

OMIM

8

Search for

Go

Clear

Limits

Preview/Index

History

Clipboard

Details

Display

default

Save

Text

Add to clipboard

1: NM_030588. Homo sapiens
DEAD...[gi:13514821]

Related Sequences, OMIM, Protein, PubMed, Taxonomy,
UniSTS, LinkOut

LOCUS NM_030588 1378 bp mRNA linear PRI 02-APR-2001

DEFINITION Homo sapiens DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 9 (RNA helicase A, nuclear DNA helicase II; leukophysin) (DDX9), transcript variant 2, mRNA.

ACCESSION NM_030588

VERSION NM_030588.1 GI:13514821

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1378)

AUTHORS Lee, C.G. and Hurwitz, J.

TITLE A new RNA helicase isolated from HeLa cells that catalytically translocates in the 3' to 5' direction

JOURNAL J. Biol. Chem. 267 (7), 4398-4407 (1992)

MEDLINE 92165790

PUBMED 1537828

REFERENCE 2 (bases 1 to 1378)

AUTHORS Lee, C.G., Zamore, P.D., Green, M.R. and Hurwitz, J.

TITLE RNA annealing activity is intrinsically associated with U2AF

JOURNAL J. Biol. Chem. 268 (18), 13472-13478 (1993)

MEDLINE 93293869

PUBMED 7685763

REFERENCE 3 (bases 1 to 1378)

AUTHORS Lee, C.G. and Hurwitz, J.

TITLE Human RNA helicase A is homologous to the maleless protein of Drosophila

JOURNAL J. Biol. Chem. 268 (22), 16822-16830 (1993)

MEDLINE 93346440

PUBMED 8344961

REFERENCE 4 (bases 1 to 1378)

AUTHORS Abdelhaleem, M.M., Hameed, S., Klassen, D. and Greenberg, A.H.

TITLE Leukophysin: an RNA helicase A-related molecule identified in cytotoxic T cell granules and vesicles

JOURNAL J. Immunol. 156 (6), 2026-2035 (1996)

MEDLINE 96310937

PUBMED 8690889

REFERENCE 5 (bases 1 to 1378)

AUTHORS Zhang, S. and Grosse, F.

TITLE Domain structure of human nuclear DNA helicase II (RNA helicase A)

JOURNAL J. Biol. Chem. 272 (17), 11487-11494 (1997)

MEDLINE 97269062

PUBMED 9111062

REFERENCE 6 (bases 1 to 1378)

AUTHORS Nakajima, T., Uchida, C., Anderson, S.F., Lee, C.G., Hurwitz, J.,

DDX9

10/510903

FIG. 8 (2/4)

Parvin, J.D. and Montminy, M.

TITLE RNA helicase A mediates association of CBP with RNA polymerase II

JOURNAL Cell 90 (6), 1107-1112 (1997)

MEDLINE 97462911

PUBMED 9323138

REFERENCE 7 (bases 1 to 1378)

AUTHORS Lee, C.G., da Costa Soares, V., Newberger, C., Manova, K., Lacy, E. and Hurwitz, J.

TITLE RNA helicase A is essential for normal gastrulation

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 95 (23), 13709-13713 (1998)

MEDLINE 99030634

PUBMED 9811865

REFERENCE 8 (bases 1 to 1378)

AUTHORS Lee, C.G., Eki, T., Okumura, K., Nogami, M., Soares, Vd., Murakami, Y., Hanaoka, F. and Hurwitz, J.

TITLE The human RNA helicase A (DDX9) gene maps to the prostate cancer susceptibility locus at chromosome band 1q25 and its pseudogene (DDX9P) to 13q22, respectively

JOURNAL Somat. Cell Mol. Genet. 25 (1), 33-39 (1999)

MEDLINE 20381755

PUBMED 10925702

COMMENT REVIEWED REFSEQ: This record has been curated by NCBI staff. The reference sequence was derived from U03643.1.
 Summary: DEAD box proteins, characterized by the conserved motif Asp-Glu-Ala-Asp (DEAD), are putative RNA helicases. They are implicated in a number of cellular processes involving alteration of RNA secondary structure such as translation initiation, nuclear and mitochondrial splicing, and ribosome and spliceosome assembly. Based on their distribution patterns, some members of this family are believed to be involved in embryogenesis, spermatogenesis, and cellular growth and division. This gene includes 2 alternatively spliced transcripts, encoding 2 different isoforms. The larger isoform is a DEAD box protein with RNA helicase activity. It may participate in melting of DNA:RNA hybrids, such as those that occur during transcription, and may play a role in X-linked gene expression. It contains 2 copies of a double-stranded RNA-binding domain, a DEXH core domain and an RGG box. The RNA-binding domains and RGG box influence and regulate RNA helicase activity. The smaller isoform is a lymphocyte granule protein. It lacks RNA-binding domains and DEXH core domain, but contains an RGG box, which may render this isoform RNA binding function.
 Transcript Variant: This variant (2) is missing a 104 nt internal fragment, in addition to 2722 nt in the 5' UTR, as compared to variant 1. It encodes the smaller isoform, which is associated with lymphocyte granules.
 COMPLETENESS: complete on the 3' end.

FEATURES

source Location/Qualifiers
 1..1378
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"
 /map="1q25"

gene 1..1378
 /gene="DDX9"
 /note="LKP; NDHII; RHA"
 /db_xref="LocusID:1660"
 /db_xref="MIM:603115"

variation 35
 /allele="A"
 /allele="G"

FIG. 8 (3/4)

10/510903

variation /db_xref="dbSNP:1049264"
 51
 /allele="A"
 /allele="G"
variation /db_xref="dbSNP:1049265"
 52
 /allele="A"
 /allele="G"
CDS /db_xref="dbSNP:1049266"
 358..1065
 /gene="DDX9"
 /note="isoform 2 is encoded by transcript variant 2; RNA
 helicase A; leukophysin; DEAD/H box-9; nuclear DNA
 helicase II; ATP-dependent RNA helicase A"
 /codon_start=1
 /db_xref="LocusID:1660"
 /db_xref="MIM:603115"
 /product="DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 9,
 isoform 2"
 /protein_id="NP_085077.1"
 /db_xref="GI:13514822"
 /translation="MKYPSPPFFVFG EKIRTR AISAKGMTLVTP LQLLLFASKK VQSDG
 QIVLVDDWIKLQISHEAAACITGLRAAMEALVVEVTKQPAIISQLDPVNERMLNMIRQ
 ISRPSAAGINLMIGSTRYGDGPRPPKMARYDNGSGYRRGGSSYSGGGYGGGYSSGGYG
 SGGYGGSSANSFRAGYGAGVGGGYRGVSRGGFRGNSGGDYRGPSGGYRGSGGFQRRGGGR
 GAYGTGYFGQGRGGGGY"
misc feature 760..1062
 /note="Arg/Gly/Ser/Tyr-rich domain; Region: RGG box"
variation 1146
 /allele="A"
 /allele="T"
variation /db_xref="dbSNP:861"
 1187
 /allele="G"
 /allele="T"
variation /db_xref="dbSNP:865"
 1236
 /allele="G"
 /allele="T"
variation /db_xref="dbSNP:860"
 1240
 /allele="A"
 /allele="T"
variation /db_xref="dbSNP:864"
 1293
 /allele="C"
 /allele="T"
variation /db_xref="dbSNP:863"
 1297
 /allele="A"
 /allele="T"
variation /db_xref="dbSNP:866"
 1318
 /allele="A"
 /allele="T"
polyA signal /db_xref="dbSNP:862"
 1362..1367
polyA site 1378
 BASE COUNT 369 a 261 c 351 g 397 t
 ORIGIN

10/510903

FIG. 8 (4/4)

1 cattgctgct gctacctgct ttccagagcc tttcatcaat gaaggaaagc ggctgggcta
61 tatccatcga aattttgctg gaaacagatt ttctgacac gtagcccttt tatcagtatt
121 ccaagcctgg gatgatgcta gaatgggtgg agaagaagca gagatacgtt tttgtgagca
181 caaaagactt aatatggcta cactaagaat gacctgggaa gccaaagtcc agctcaaaga
241 gatttttgatt aattctgggt ttccagaaga ttgtttgttg acacaagtgt ttactaacac
301 tggaccagat aataatttg atgttggttat ctccctcctg gcctttgtag ccaagacatg
361 aagtacccat ctcccttctt tgtatttggg gaaaagattc gaactcgagc catctctgct
421 aaaggcatga ctttagtcac cccctgcag ttgcttctct ttgcctccaa gaaagtccaa
481 tctgatgggc agatttgctg tgtagatgac tggattaaac tgcaaatatc tcatgaagct
541 gctgcctgta tcaactggtc ccgggcagcc atggaggctt tggttggtga agtaaccaa
601 caacctgcta tcatcagcca gttggacccc gtaaatgaac gtatgctgaa catgatccgt
661 cagatctcta gaccctcagc tgctggtatc aaccttatga ttggcagtag acggtatgga
721 gatgggtccac gtccctccaa gatggcccga tacgacaatg gaagcggata tagaagggga
781 ggttctagtt acagtgggtg aggctatggc ggtggctata gcagtggagg ctatggtagc
841 ggaggctatg gtggcagcgc caactccttt cgggcaggat atggtgcagg tgttggtgga
901 ggctatagag gagtttcccg aggtggcttt agaggcaact ctggaggaga ctacagaggg
961 cctagtggag gctacagagg atctggggga ttccagcgag gaggtggtag gggggcctat
1021 ggaactggct actttggaca gggaagagga ggtggcggct attaaaactt ggttatgtca
1081 gttcctgtgt gtagacagta aggaaaaaaa ggcattgctat gtgttacgtg ttttttccag
1141 tatgtttatt tgccaccaa aagtaaattgc attttcaccc attctgtggt tcattgtagt
1201 ttaaggaaac caagcatata gatgcattag tgattttgtt tatattatgt aaaatataac
1261 gatctcttaa aaataccaca gtttgtattt tttctttaag gagtaaagat ttgcctttaa
1321 ataacttggg attttctctg ctttcgttta atacaataga aaataaagta ttacaccg

//

Revised: October 24, 2001.

[Disclaimer](#) | [Write to the Help Desk](#)
[NCBI](#) | [NLM](#) | [NIH](#)

10/510903

FIG. 9 (1/6)



Nucleotide

☐ PubMed ☐ Nucleotide ☐ Protein ☐ Genome ☐ Structure ☐ PopSet ☐ Taxonomy ☐ OMIM ☐ B

Search for

1: NM_000875. Homo sapiens
insu...[gi:11068002]

Related Sequences, OMIM, Protein, PubMed, Taxonomy,
UniSTS, LinkOut

LOCUS NM_000875 4989 bp mRNA linear PRI 01-NOV-2000
 DEFINITION Homo sapiens insulin-like growth factor 1 receptor (IGF1R), mRNA.
 ACCESSION NM_000875
 VERSION NM_000875.2 GI:11068002
 KEYWORDS IGF1-R
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 4989)
 AUTHORS Flier JS, Usher P and Moses AC.
 TITLE Monoclonal antibody to the type I insulin-like growth factor
 (IGF-I) receptor blocks IGF-I receptor-mediated DNA synthesis:
 clarification of the mitogenic mechanisms of IGF-I and insulin in
 human skin fibroblasts
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 83 (3), 664-668 (1986)
 MEDLINE 86121000
 PUBMED 3003744

REFERENCE 2 (bases 1 to 4989)
 AUTHORS Francke U, Yang-Feng TL, Brissenden JE and Ullrich A.
 TITLE Chromosomal mapping of genes involved in growth control
 JOURNAL Cold Spring Harb. Symp. Quant. Biol. 51 Pt 2, 855-866 (1986)
 MEDLINE 87217109
 PUBMED 3107886

REFERENCE 3 (bases 1 to 4989)
 AUTHORS Ullrich, A., Gray, A., Tam, A.W., Yang-Feng, T., Tsubokawa, M.,
 Collins, C., Henzel, W., Bon, T.L., Kathuria, S., Chen, E., Jakobs, S.,
 Francke, U., Ramachandran, J. and Fujita-Yamaguchi, Y.
 TITLE Insulin-like growth factor I receptor primary structure: comparison
 with insulin receptor suggests structural determinants that define
 functional specificity
 JOURNAL EMBO J. 5 (10), 2503-2512 (1986)
 MEDLINE 87053815

REFERENCE 4 (bases 1 to 4989)
 AUTHORS Cooke DW, Bankert LA, Roberts CT Jr, LeRoith D and Casella SJ.
 TITLE Analysis of the human type I insulin-like growth factor receptor
 promoter region
 JOURNAL Biochem. Biophys. Res. Commun. 177 (3), 1113-1120 (1991)
 MEDLINE 91282751
 PUBMED 1711844

REFERENCE 5 (bases 1 to 4989)
 AUTHORS Abbott AM, Bueno R, Pedrini MT, Murray JM and Smith RJ.
 TITLE Insulin-like growth factor I receptor gene structure
 JOURNAL J. Biol. Chem. 267 (15), 10759-10763 (1992)
 MEDLINE 92268129
 PUBMED 1316909

FIG. 9 (2/6)

10/510903

REFERENCE 6 (bases 1 to 4989)
 AUTHORS Werner H, Karnieli E, Rauscher FJ and LeRoith D.
 TITLE Wild-type and mutant p53 differentially regulate transcription of the insulin-like growth factor I receptor gene
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 93 (16), 8318-8323 (1996)
 MEDLINE 96323219
 PUBMED 8710868

REFERENCE 7 (bases 1 to 4989)
 AUTHORS Grant ES, Ross MB, Ballard S, Naylor A and Habib FK.
 TITLE The insulin-like growth factor type I receptor stimulates growth and suppresses apoptosis in prostatic stromal cells
 JOURNAL J. Clin. Endocrinol. Metab. 83 (9), 3252-3257 (1998)
 MEDLINE 98417960
 PUBMED 9745438

COMMENT REVIEWED REFSEQ: This record has been curated by NCBI staff. The reference sequence was derived from X04434.1, M69229.1.
 On Nov 1, 2000 this sequence version replaced gi:4557664.
 Summary: This receptor binds insulin-like growth factor with a high affinity. It has tyrosine kinase activity. The insulin-like growth factor I receptor plays a critical role in transformation events. Cleavage of the precursor generates alpha and beta subunits. It is highly overexpressed in most malignant tissues where it functions as an anti-apoptotic agent by enhancing cell survival.

FEATURES
 source Location/Qualifiers
 1..4989
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="15"
 /map="15q25-q26"
 /clone="(lambda) IGF-1-R.85, (lambda) IGF-1-R.76"
 /tissue_type="placenta"
 /clone_lib="(lambda)gt10"

gene
 1..4989
 /gene="IGF1R"
 /note="JTK13"
 /db_xref="LocusID:3480"
 /db_xref="MIM:147370"

CDS
 46..4149
 /gene="IGF1R"
 /EC_number="2.7.1.112"
 /codon_start=1
 /db_xref="LocusID:3480"
 /db_xref="MIM:147370"
 /product="insulin-like growth factor 1 receptor precursor"
 /protein_id="NP_000866.1"
 /db_xref="GI:4557665"
 /translation="MKSGSGGGSPTSLWGLLFLSAALSLWPTSGEICGPGIDIRNDYQ
 QLKRENTVIEGYLHILLISKAEDYRSYRFPKLTVITEYLLLFRVAGLESIGDLFPN
 LTVIRGWKLFYNYALVIFEMTNLKDIGLYNLRNITRGAIRIEKNADLCYLSTVDWSLI
 LDAVSNNYIVGNKPPKECGDLCPGTMEEEKPMCEKTTINNEYNYRCWTTNRCQKMCPT
 CGKRACTENNECCHPECLGSCSAPDNDTACVACRHHYVYAGVCVPACPPNTYRFEGRWC
 VDRDFCANILSAESSDSEGFVIHDGECMQECPSGFIRNGSQSMYCIPCEGPCPKVCEE
 EKTKTIDSVTSAQMLQGCTIFKGNLLINIRGNINIAELENFMGLIEVVTGYVKIRH
 SHALVSLFLKNLRLILGEEQLEGNYSFYVLDNQNLQQLWDWDHRLNLTIKAGKMYFAF
 NPKLCVSEIYRMEEVTGTGKRQSKGDINTRNGERASCESDVLHFTSTTTSKNRIIT
 WHRYRPPDYRDLISFTVYVYKEAPFKNVTEYDGDACGSNSWNMVDVDLPPNKDVEPGI
 LLHGLKPWTQYAVYVKAFTLTMVENDHIRGAKSEILYIRTNASVPSIPLDVLASNSS
 SQLIVKWNPPSLPNGNLSYIVRWQRQPDGYLYRHNYCSKDKIPIRKYADGTIDIEE
 VTENPKTEVCGGEKGPCCACPKTEAEKQAEKEEAERYKVFENFLHNSIFVPRPERKRR
 DVMQVANTTMSSRSRNTTAADTYNITDPEELETEYPPFFESRVDNKERTVISNLRPFTL

FIG. 9 (3/6)

10/510903

YRIDIHSCNHEAEKLGCSASN FVFARTMPAEGADDIPGPVTWEPRPENSTFLKWPEPE
 NPNGLILMYEIKYGSQVEDQRECVSRQEYRKYGAKLNRLNPGNYTARIQATSLSGNG
 SWTDPVFFYVQAKTGYENFIHLIIALPVAVLLIVGGLVIMLYVFHRKRNSRLGNGVL
 YASVNPEYFSAADVVPDEWEVAREKITMSRELGGQSGFGMVYEGVAKGVVKDEPETRV
 AIKTVNEAASMRERIEFLNEASVMKEFNCHHVRLLGVVSSQGQPTLVIMELMTRGDLK
 SYLRSLRPEMENNPFVLAPPSLSKMIQMAGEIADGMAYLNANKFVHRDLAARNCMVAED
 FTVKIGDFGMTRDIYETDYRKGGKGLLPVRWMSPELKDGVFTTYSVDVWSFGVVWE
 IATLAEQPYQGLSNEQVLR FVMEGGLLDKPDNCPDMLFELMRMCWQYNPKMRPSFLEI
 ISSIKEEMEPGFREVVSFYIYSEENKLEPEEELDLEPENMESVPLDPSASSSSSLPLPDRH
 SGHKAENGPGPGVLVLRASFDERQPYAHMNGGRKNERALPLPQSSTC"

sig_peptide 46..123
mat_peptide 121..4134
 /product="IGF-I receptor"
misc_feature 122..2251
 /note="alpha-subunit (AA 1 - 710)"
misc_feature 182..190
 /note="pot.N-linked glycosylation site (AA 21 - 23)"
misc_feature 196..561
 /note="Recep_L_domain; Region: Receptor L domain"
misc_feature 335..343
 /note="pot.N-linked glycostlation site (AA 72 - 74)"
misc_feature 434..442
 /note="pot.N-linked glycostlation site (AA 105 - 107)"
misc_feature 568..1044
 /note="Furin-like; Region: Furin-like cysteine rich region"
misc_feature 724..852
 /note="FU; Region: Furin-like repeats"
misc_feature 761..769
 /note="pot.N-linked glycostlation site (AA 214 - 216)"
misc_feature 971..979
 /note="pot.N-linked glycostlation site (AA 284 - 286)"
misc_feature 1162..1410
 /note="Recep_L_domain; Region: Receptor L domain"
misc_feature 1280..1288
 /note="pot.N-linked glycostlation site (AA 387 - 389)"
misc_feature 1343..1351
 /note="pot.N-linked glycosylation site (AA 408 - 410)"
misc_feature 1631..1639
 /note="pot.N-linked glycostlation site (AA 504 - 506)"
variation 1731
 /allele="A"
 /allele="G"
 /db_xref="dbSNP:2228531"
misc_feature 1850..1858
 /note="pot.N-linked glycosylation site (AA 577 - 579)"
misc_feature 1895..1903
 /note="pot.N-linked glycosylation site (AA 592 - 594)"
misc_feature 1949..1957
 /note="pot.N-linked glycosylation site (AA 610 - 612)"
misc_feature 2240..2251
 /note="putative proreceptor processing site (AA 707 - 710)"
misc_feature 2252..4132
 /note="beta-subunit (AA 711 - 1337)"
misc_feature 2270..2278
 /note="pot.N-linked glycosylation site (AA 717 - 719)"
misc_feature 2297..2305
 /note="pot.N-linked glycosylation site (AA 726 - 728)"
misc_feature 2321..2329

10/510903

FIG. 9 (4/6)

misc feature /note="pot.N-linked glycosylation site (AA 734 - 736)"
 2548..2796
misc feature /note="fn3; Region: Fibronectin type III domain"
 2729..2737
misc feature /note="pot.N-linked glycosylation site (AA 870 - 872)"
 2768..2776
misc feature /note="pot.N-linked glycosylation site (AA 883 - 885)"
 2836..2910
misc feature /note="transmembrane region (AA 906 - 929);
 transmembrane-region site"
 2918..2926
misc feature /note="pot.N-linked glycosylation site (AA 933 - 935)"
 3040..3834
misc feature /note="pkinase; Region: Eukaryotic protein kinase domain"
 3040..3843
misc feature /note="TyrKc; Region: Tyrosine kinase, catalytic domain"
 3047..3049
misc feature /note="pot.ATP binding site (AA 976)"
 3052..3807
misc feature /note="S_TKc; Region: Serine/Threonine protein kinases,
 catalytic domain"
 3053..3055
misc feature /note="pot.ATP binding site (AA 978)"
 3062..3064
misc feature /note="pot.ATP binding site (AA 981)"
 3128..3130
variation /note="pot.ATP binding site (AA 1003)"
 4267
 /allele="A"
 /allele="T"
 /db_xref="dbSNP:1065304"
variation 4268
 /allele="A"
 /allele="T"
 /db_xref="dbSNP:1065305"

BASE COUNT 1216 a 1371 c 1320 g 1082 t
 ORIGIN

```

1 tttttttttt ttttgagaaa gggaatttca tcccaaataa aaggaatgaa gtctggetcc
61 ggaggagggt ccccgacctc gctgtggggg ctctgttttc tctccgccgc gctctcgctc
121 tggccgacga gtggagaaat ctgcggggcca ggcacgcaca tccgcaacga ctatcagcag
181 ctgaagcgcc tggagaactg cacggtgatc gagggctacc tccacatcct gctcatctcc
241 aaggccgagg actaccgcag ctaccgcttc cccaagctca cggtcattac cgagtacttg
301 ctgctgttcc gagtggctgg cctcgagagc ctccggagacc tcttcccca cctcacggtc
361 atccgcggct ggaaactctt ctacaactac gccctgggtca tcttcgagat gaccaatctc
421 aaggatattg ggctttacaa cctgagggaac attactcggg gggccatcag gattgagaaa
481 aatgctgacc tctgttacct ctccactgtg gactgggtccc tgatcctgga tgcggtgtcc
541 aataactaca ttgtggggaa taagccccc aaggaatgtg gggacctgtg tccagggacc
601 atggaggaga agccgatgtg tgagaagacc accatcaaca atgagtacaa ctaccgctgc
661 tggaccacaa accgctgcca gaaaatgtgc ccaagcacgt gtgggaagcg ggcgtgcacc
721 gagaacaatg agtgtgcca ccccgagtgc ctgggcagct gcagcgcgcc tgacaacgac
781 acggcctgtg tagcttgccg ccactactac tatgccggtg tctgtgtgcc tgccctgccg
841 cccaacacct acaggtttga gggctggcgc tgtgtggacc gtgacttctg cgccaacatc
901 ctcagcgccg agagcagcga ctccgagggg tttgtgatcc acgacggcga gtgcattgcag
961 gagtgtccct cgggcttcat ccgcaacggc agccagagca tgtactgcat cccttgtaga
1021 ggtccttgcc cgaaggtctg tgaggaagaa aagaaaacaa agaccattga ttctgttact
1081 tctgctcaga tgctccaagg atgcaccatc ttcaagggca atttgctcat taacatccga
1141 cgggggaata acattgcttc agagctggag aacttcattg ggctcatcga ggtggtgacg
1201 ggctacgtga agatccgcca ttctcatgcc ttggtctcct tgtccttcct aaaaaacctt
1261 cgctcatcc taggagagga gcagctagaa gggaattact ccttctacgt cctcgacaac
1321 cagaacttg cagcaactgtg ggactgggac caccgcaacc tgaccatcaa agcagggaaa
  
```

FIG. 9 (5/6)

10/510903

```

1381 atgtactttg ctttcaatcc caaattatgt gtttccgaaa tttaccgcat ggaggaaagtg
1441 acgggggacta aagggcgcca aagcaaaggg gacataaaca ccaggaacaa cggggagaga
1501 gcctcctgtg aaagtgacgt cctgcatttc acctccacca ccacgtcgaa gaatcgcatc
1561 atcataacct ggcaccggta ccggcccccct gactacaggg atctcatcag cttcaccgtt
1621 tactacaagg aagcaccctt taagaatgtc acagagtatg atgggcagga tgcctgcggc
1681 tccaacagct ggaacatggt ggacgtggac ctcccgcaca acaaggacgt ggagcccggc
1741 atcttactac atgggctgaa gccctggact cagtacgccg tttacgtcaa ggctgtgacc
1801 ctcaccatgg tggagaacga ccatatccgt ggggccaaga gtgagatctt gtacattcgc
1861 accaatgctt cagttccttc cattcccttg gacgttcttt cagcatcgaa ctctcttct
1921 cagttaatcg tgaagtggaa cctcctctct ctgcccacag gcaacctgag ttactacatt
1981 gtgcgctggc agcggcagcc tcaggacggc tacctttacc ggcacaatta ctgctccaaa
2041 gacaaaatcc ccatcaggaa gtatgccgac ggcaccatcg acattgagga ggtcacagag
2101 aacccaaga ctgaggtgtg tgggtggggag aaagggcctt gctgcgcctg ccccaaaact
2161 gaagccgaga agcaggccga gaaggaggag gctgaatacc gcaaagtctt tgagaatttc
2221 ctgcacaact ccatcttctg gcccagacct gaaaggaagc ggagagatgt catgcaagtg
2281 gccaacacca ccatgtccag ccgaagcagg aacaccacgg ccgcagacac ctacaacatc
2341 accgaccgg aagagctgga gacagatcac cctttctttg agagcagagt ggataaaga
2401 gagagaactg tcaattctaa ccttcggcct ttcacattgt accgcatcga tatccacagc
2461 tgcaaccacg aggtgagaa gctgggctgc agcgcctcca acttcgtctt tgcaaggact
2521 atgcccgcag aaggagcaga tgacattcct gggccagtga cctgggagcc aaggcctgaa
2581 aactccatct ttttaaagtg gccggaacct gagaatccca atggattgat tctaattgat
2641 gaaataaaat acggatcaca agttgaggat cagcgagaat gtgtgtccag acaggaatac
2701 aggaagtatg gaggggccaa gctaaaccgg ctaaaccggg ggaactacac agcccggatt
2761 caggccacat ctctctctgg gaatgggtcg tggacagatc ctgtgttctt ctatgtccag
2821 gccaaaacag gatatgaaaa cttcatccat ctgatcatcg ctctgcccgt cgctgtcctg
2881 ttgatcgtgg gaggttgggt gattatgctg tacgtcttcc atagaaagag aaataacagc
2941 aggtcgtggg atggagtgc gtatgcctct gtgaaccggg agtacttcag cgctgtgat
3001 gtgtacgttc ctgatgagt ggaggtggct cgggagaaga tcaccatgag ccgggaactt
3061 gggcaggggt cgtttgggat ggtctatgaa ggagttgcca aggtgtggt gaaagatgaa
3121 cctgaaacca gagtggccat taaaacagtg aacgaggccg caagcatgcg tgagaggatt
3181 gagtttctca acgaagcttc tgtgatgaag gagttcaatt gtcaccatgt ggtgcgattg
3241 ctgggtgtgg tgtcccaagg ccagccaaca ctggtcatca tggaaactgat gacacggggc
3301 gatctcaaaa gttatctccg gtctctgagg ccagaaatgg agaataatcc agtcctagca
3361 cctccaagc tgagcaagat gattcagatg gccggagaga ttgcagacgg catggcatac
3421 ctcaacgcca ataagttcgt ccacagagac cttgctgcc ggaattgcag ggtagccgaa
3481 gatttcacag tcaaaatcgg agattttgggt atgacgcgag atatctatga gacagactat
3541 taccggaaag gaggcaaagg gctgctgccc gtgcgctgga tgtctcctga gtccctcaag
3601 gatggagtct tcaccactta ctcgacgtc tggctcttcg gggctgctct ctgggagatc
3661 gccacactgg ccgagcagcc ctaccagggc ttgtccaacg agcaagtcct tcgcttcgtc
3721 atggagggcg gccttctgga caagccagac aactgtcctg acatgctgtt tgaactgatg
3781 cgcatgtgct ggcagtataa cccaagatg aggccttctt tcctggagat catcagcagc
3841 atcaaagagg agatggagcc tggcttccgg gaggtctctt tctactacag cgaggagaac
3901 aagctgcccg agccggagga gctggaccta gagccagaga acatggagag cgtcccctg
3961 gaccctcgg cctcctcgtc ctcccgcga ctggccgaca gacactcag acacaaggcc
4021 gagaacggcc ccggccctgg ggtgctggtc ctccgcgcca gcttcgacga gagacagcct
4081 tacgcccaca tgaacggggg ccgcaagaac gagcgggcct tgcgctgcc ccagtcttcg
4141 acctgctgat ccttgatcc tgaatctgtg caaacagtaa cgtgtgcgca cgcgcagcgg
4201 ggtggggggg gagagagagt ttaacaatc cattcacaag cctcctgtac ctcagtggat
4261 cttcagttct gcccttgctg ccgcggggag acagcttctc tgcagtaaaa cacatttggg
4321 atgttccctt tttcaatatg caagcagctt ttatttccct gcccaaacc ttaactgaca
4381 tgggccttta agaaccttaa tgacaacact taatagcaac agagcacttg agaaccagtc
4441 tctcactct gtccctgtcc tccctgttc tcccttctc tctcctctc gcttcataac
4501 ggaaaataa ttgccacaag tccagctggg aagccctttt tatcagtttg aggaagtggc
4561 tgtccctgtg gccccatcca accactgtac acaccgcct gacaccgtgg gtcattacaa
4621 aaaaacacgt ggagatggaa atttttacct ttatctttca cctttctagg gacatgaaat
4681 ttacaaaggg ccatcgttca tccaaggctg ttaccatttt aacgctgcct aattttgcca
4741 aaatcctgaa ctttctccct catcgccccg gcgctgattc ctctgttccg gaggcagggg
4801 tgagcatggc agctggttgc tccatttgag agacacgctg gcgacacact ccgtccatcc
4861 gactgcccct gctgtgctgc tcaaggccac aggcacacag gtctcattgc ttctgactag
4921 attattattt gggggaactg gacacaatag gtctttctct cagtgaagggt ggggagaagc

```

10/510903

FIG. 9 (6/6)

4981 tgaaccggc

//

Revised: October 24, 2001.

Disclaimer | Write to the Help Desk
NCBI | NLM | NIH

10/510903

FIG. 10 (1/4)



Nucleotide

PubMed

Nucleotide

Protein

Genome

Structure

PopSet

Taxonomy

OMIM

Search for
 Limits

 1: NM_003349. Homo sapiens
 ubiq...[gi:15718757]

 Related Sequences, OMIM, Protein, PubMed, Taxonomy,
 UniSTS, LinkOut

LOCUS NM_003349 2394 bp mRNA linear PRI 21-SEP-2001
 DEFINITION Homo sapiens ubiquitin-conjugating enzyme E2 variant 1 (UBE2V1),
 transcript variant 2, mRNA.
 ACCESSION NM_003349
 VERSION NM_003349.3 GI:15718757 **UBE2 V1**
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 2394)
 AUTHORS Rothofsky, M.L. and Lin, S.L.
 TITLE CROC-1 encodes a protein which mediates transcriptional activation
 of the human FOS promoter
 JOURNAL Gene 195 (2), 141-149 (1997)
 MEDLINE 97449289
 PUBMED 9305758
 REFERENCE 2 (bases 1 to 2394)
 AUTHORS Sancho, E., Vila, M.R., Sanchez-Pulido, L., Lozano, J.J., Paciucci,
 R., Nadal, M., Fox, M., Harvey, C., Bercovich, B., Loukili, N.,
 Ciechanover, A., Lin, S.L., Sanz, F., Estivill, X., Valencia, A.
 and Thomson, T.M.
 TITLE Role of UEV-1, an inactive variant of the E2 ubiquitin-conjugating
 enzymes, in in vitro differentiation and cell cycle behavior of
 HT-29-M6 intestinal mucosecretory cells
 JOURNAL Mol. Cell. Biol. 18 (1), 576-589 (1998)
 MEDLINE 98078713
 PUBMED 9418904
 REFERENCE 3 (bases 1 to 2394)
 AUTHORS Ma, L., Broomfield, S., Lavery, C., Lin, S.L., Xiao, W. and Bacchetti, S.
 TITLE Up-regulation of CIR1/CROC1 expression upon cell immortalization
 and in tumor-derived human cell lines
 JOURNAL Oncogene 17 (10), 1321-1326 (1998)
 MEDLINE 98442973
 PUBMED 9771976
 REFERENCE 4 (bases 1 to 2394)
 AUTHORS Hofmann, R.M. and Pickart, C.M.
 TITLE Noncanonical MMS2-encoded ubiquitin-conjugating enzyme functions in
 assembly of novel polyubiquitin chains for DNA repair
 JOURNAL Cell 96 (5), 645-653 (1999)
 MEDLINE 99189750
 PUBMED 10089880
 REFERENCE 5 (bases 1 to 2394)
 AUTHORS Deng, L., Wang, C., Spencer, E., Yang, L., Braun, A., You, J.,
 Slaughter, C., Pickart, C. and Chen, Z.J.
 TITLE Activation of the IkappaB kinase complex by TRAF6 requires a

10/510903

FIG. 10 (2/4)

dimeric ubiquitin-conjugating enzyme complex and a unique polyubiquitin chain

JOURNAL Cell 103 (2), 351-361 (2000)

MEDLINE 20509589

PUBMED 11057907

REFERENCE 6 (bases 1 to 2394)

AUTHORS Thomson,T.M., Lozano,J.J., Loukili,N., Carrio,R., Serras,F., Cormand,B., Valeri,M., Diaz,V.M., Abril,J., Burset,M., Merino,J., Macaya,A., Corominas,M. and Guigo,R.

TITLE Fusion of the human gene for the polyubiquitination coeffectector UEV1 with Kua, a newly identified gene

JOURNAL Genome Res. 10 (11), 1743-1756 (2000)

MEDLINE 20530912

PUBMED 11076860

COMMENT REVIEWED REFSEQ: This record has been curated by NCBI staff. The reference sequence was derived from U39361.1, AL110132.1. On Sep 21, 2001 this sequence version replaced gi:12025659. Summary: Ubiquitin-conjugating enzyme E2 variant proteins constitute a distinct subfamily within the E2 protein family. They have sequence similarity to other ubiquitin-conjugating enzymes but lack the conserved cysteine residue that is critical for the catalytic activity of E2s. The protein encoded by this gene is located in the nucleus and can cause transcriptional activation of the human FOS proto-oncogene. It is thought to be involved in the control of differentiation by altering cell cycle behaviour. Multiple alternatively spliced transcripts encoding different isoforms have been described for this gene. Transcript Variant: This variant (2) encodes the longest isoform (b) of this protein. COMPLETENESS: complete on the 3' end.

FEATURES

source Location/Qualifiers

1..2394

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="20"

/map="20q13.2"

gene 1..2394

/gene="UBE2V1"

/note="CIR1; UEV-1; UEV1; UEV1A; CROC-1; CROC1"

/db_xref="LocusID:7335"

/db_xref="MIM:602995"

CDS 70..735

/gene="UBE2V1"

/note="isoform b is encoded by transcript variant 2; DNA-binding protein"

/codon_start=1

/db_xref="LocusID:7335"

/db_xref="MIM:602995"

/product="ubiquitin-conjugating enzyme E2 variant 1, isoform b"

/protein_id="NP_003340.1"

/db_xref="GI:4507795"

/translation="MAYKFRTHSPEALEQLYPWECFVFCLIIIFGTFTNQIHKWSHTYFGLPRWVTLLQDWHVILPRKHHRIHHVSPHETYFCITTGVKVPNRNRLLEELEEGQKGVGDGTVSWGLEDDEDMTLTRWTGMIIGPPRTIYENRIYSLKIECGPKYPEAPPFVRFVTKINMNGVNSSNGVVDPAISVLAKWQNSYSIKVVLQELRRLMMSKENMKLPQPPEGQCYSN"

misc_feature 334..714

/note="UBCc; Region: Ubiquitin-conjugating enzyme E2, catalytic domain homologues"

FIG. 10 (3/4)

10/510903

misc feature 337..555
/note="UQ_con; Region: Ubiquitin-conjugating enzyme.
Proteins destined for proteasome-mediated degradation may
be ubiquitinated. Ubiquitination follows conjugation of
ubiquitin to a conserved cysteine residue of UBC
homologues. TSG101 is one of several UBC homologues that
lacks this active site cysteine"

misc feature 643..714
/note="Region: DNA-binding domain"

variation 1117
/allele="C"
/allele="T"
/db_xref="dbSNP:8585"

variation 1257
/allele="C"
/allele="T"
/db_xref="dbSNP:1049679"

variation complement(1968)
/allele="A"
/allele="G"
/db_xref="dbSNP:2733"

variation 2017
/allele="A"
/allele="C"
/db_xref="dbSNP:15218"

polyA signal 2112..2117
polyA site 2135
/evidence=experimental

variation complement(2179)
/allele="G"
/allele="T"
/db_xref="dbSNP:2664563"

variation 2249
/note="WARNING: map location ambiguous"
/allele="A"
/allele="T"
/db_xref="dbSNP:1049871"

variation complement(2259)
/allele="A"
/allele="G"
/db_xref="dbSNP:2664532"

polyA signal 2350..2355
polyA site 2373
/evidence=experimental

BASE COUNT 658 a 605 c 481 g 650 t
ORIGIN

```

1 ttcacacggc acgacttcat cgagaccaac ggggacaact gcctgggtgac actgctgccg
61 ctgctaaaca tggcctacaa gttccgcacc cacagccctg aagccctgga gcagctatac
121 ccctgggagt gcttcgtctt ctgcctgata atcttcggca ccttcaccaa ccagatccac
181 aagtggctgc acacgtactt tgggctgcc aagctgggtc cctcctgca ggactggcat
241 gtcattcctgc cagctaaaca ccatcgatc caccacgtct caccaccaga gacctacttc
301 tgcattacca caggagtaaa agtccctcgc aatttccgac tggtggaaga actcgaagaa
361 ggccagaaaag gagtaggaga tggcacagtt agctgggggtc tagaagatga cgaagacatg
421 acacttacaa gatggacagg gatgataatt gggcctccaa gaacaattta tgaaaaccga
481 atatacagcc ttaaaataga atgtggacct aaatacccag aagcaccgcc ctttgtaaga
541 tttgtaacaa aaattaatat gaattggagta aatagttcta atggagtggg ggaccaaga
601 gccatatcag tgctagcaaa atggcagaat tcatatagca tcaaagttgt cctgcaagag
661 ctctggcgcc taatgatgtc taaagaaaat atgaaactcc ctccagccgc cgaaggacag
721 tgttacagca attaatcaaa aagaaaaacc acaggccctt ccccttcccc ccaattcgat
781 ttaatcagtc ttcattttcc acagtagtaa attttctaga tacgtcttgt agacctcaaa

```

10/510903

FIG. 10 (4/4)

841 gtaccggaaa ggaagctccc attcaaagga aatttatctt aagatactgt aaatgatact
901 aattttttgt ccatttgaaa tatataagtt gtgctataac aaatcatcct gtcaagtgtgta
961 accactgtcc acgtagttga acttctggga tcaagaaagt ctattttaa atgattcccat
1021 cataactggt ggggcacatc taactcaact gtgaaaagac acatcacaca atcaccttgc
1081 tgctgattac acggcctggg gtctctgcct tctcccttta cctcccgc tcccaccctc
1141 cctgcaacaa cagccctcta gcctgggggg cttgttagag tagatgtgaa ggtttcaggt
1201 cgcagcctgt gggactactg ctaggtgtgt ggggtgttgc gctgcaccc ctggttccct
1261 taagtcttaa gtgatgcccc ttccaaacca tcatcctgtc cccacgctcc tccactcccg
1321 cccttgcccg aagcatagat tgtaaccctt cactccct ctgagattgg cttcgggtgag
1381 gaattcaggg ctttcccat atcttctctc cccccacctt tatcgagggg tgctgctttt
1441 tctccctcct cctcaagtcc ctttttgac cgtcaccacc caacacctt catgacactt
1501 ccttgctttg gccagaagcc atcaggtgaa gttggaaaga gcctctgacc tcccttggtt
1561 agttttggaa ccatactcac tcaactctca ccagcctggg aaatgaatat tgggtcctca
1621 gccctgccac cctctgctgt catcagctga tgcattgttt ttagctcagg ttttgataag
1681 gtgaaaagaa tagtcaccag ggttactcag acctgccagc tctcgagtc cttggtggtt
1741 gaacttgag aaagaccgca tgaagatact tgtaagcaca catgatccct ctgaattggt
1801 ttactttcct gtaactgctt ttgcttttaa aaattgaaga agtttgaat ctgacaactg gaacaaaaag
1861 tttggtcatc cttgcaatcc attgggtct ctttgggtt ggtgctgctg ctgcttccca agatcctcag
1921 aaccttgaat ccggtgcatg ccttgggtt ggtgctgctg ctgcttccca agatcctcag
1981 cagggattaa gaaggaaccc ggtgtgcaca gcagatcccc gaaattggtg ggcttgacct
2041 cctggcaaat tgctgcgtct ttccacttgc tgttcaggac cactaaatgc tgaaatgtgg
2101 atgcataccg aaataaaagc aattcattgt gtactaaagg tttttttttt ttttttaatt
2161 tagtatttgt gtaaaaccac cttttgaagc agcaactatc aagtctgaaa agcaattgat
2221 gtttccatta atctttttct ggggggaaaa ccttagttct aaggatttaa catcctgtaa
2281 gtgaagttta acataacagt attccataag cagccttttt attgtcagac cattgcctga
2341 ttttaataata ataaaaaaaa agtgtgcgtt aataaaaaaaaa aaaaaaaaaa aaaa

//

Revised: October 24, 2001.

Disclaimer | Write to the Help Desk
NCBI | NLM | NIH

10/510903



FIG. 11 (1/3)



Nucleotide

Search for

Display

1: NM_000689. Homo sapiens
alde...[gi:4502030]

Related Sequences, OMIM, Protein, PubMed, Taxonomy,
LinkOut

LOCUS NM_000689 1506 bp mRNA linear PRI 31-OCT-2000
 DEFINITION Homo sapiens aldehyde dehydrogenase 1, soluble (ALDH1), mRNA.
 ACCESSION NM_000689
 VERSION NM_000689.1 GI:4502030
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens

ALDEHYDE DEHYDROGENASE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1506)
 AUTHORS Hsu LC, Tani K, Fujiyoshi T, Kurachi K and Yoshida A.
 TITLE Cloning of cDNAs for human aldehyde dehydrogenases 1 and 2
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 82 (11), 3771-3775 (1985)
 MEDLINE 85216574
 PUBMED 2987944

REFERENCE 2 (bases 1 to 1506)
 AUTHORS Raghunathan L, Hsu LC, Klisak I, Sparkes RS, Yoshida A and Mohandas T.
 TITLE Regional localization of the human genes for aldehyde dehydrogenase-1 and aldehyde dehydrogenase-2
 JOURNAL Genomics 2 (3), 267-269 (1988)
 MEDLINE 88284707
 PUBMED 3397064

REFERENCE 3 (bases 1 to 1506)
 AUTHORS Hsu LC, Chang WC and Yoshida A.
 TITLE Genomic structure of the human cytosolic aldehyde dehydrogenase gene
 JOURNAL Genomics 5 (4), 857-865 (1989)
 MEDLINE 90077427
 PUBMED 2591967

REFERENCE 4 (bases 1 to 1506)
 AUTHORS Pereira F, Rosenmann E, Nylen E, Kaufman M, Pinsky L and Wrogemann K.
 TITLE The 56 kDa androgen binding protein is an aldehyde dehydrogenase
 JOURNAL Biochem. Biophys. Res. Commun. 175 (3), 831-838 (1991)
 MEDLINE 91222190
 PUBMED 1709013

REFERENCE 5 (bases 1 to 1506)
 AUTHORS Zheng, C.F., Wang, T.T. and Weiner, H.
 TITLE Cloning and expression of the full-length cDNAs encoding human liver class 1 and class 2 aldehyde dehydrogenase
 JOURNAL Alcohol. Clin. Exp. Res. 17 (4), 828-831 (1993)
 MEDLINE 94027752

REFERENCE 6 (bases 1 to 1506)
 AUTHORS Kathmann, E.C. and Lipsky, J.J.
 TITLE Cloning of a cDNA encoding a constitutively expressed rat liver

FIG. 11 (2/3)

10/510903

cytosolic aldehyde dehydrogenase
JOURNAL Biochem. Biophys. Res. Commun. 236 (2), 527-531 (1997)
MEDLINE 97382470
COMMENT PROVISIONAL REFSEQ: This record has not yet been subject to final NCBI review. The reference sequence was derived from AF003341.1.

FEATURES

source Location/Qualifiers
 1..1506
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="9"
 /map="9q21"
 /tissue_type="liver"

gene 1..1506
 /gene="ALDH1"
 /note="PUMB1"
 /db_xref="LocusID:216"
 /db_xref="MIM:100640"

CDS 1..1506
 /gene="ALDH1"
 /EC_number="1.2.1.3"
 /note="cytosolic protein; class 1"
 /codon_start=1
 /db_xref="LocusID:216"
 /db_xref="MIM:100640"
 /product="aldehyde dehydrogenase 1, soluble"
 /protein_id="NP_000680.1"
 /db_xref="GI:4502031"
 /translation="MSSSGTPDLPVLLTDLKIQYTKIFINNEWHDSVSGKKFPVFNPA
 TEEELCQVEEGDKEDVDKAVKAARQAFQIGSPWRTMDASERGRLLYKLADLIERDLL
 LATMESMNGGKLYSNAYLSDLAGCIKTLRYCAGWADKIQGRTIPIDGNFFTYTRHEPI
 GVCQGIIIPWNFPLVMLIWKIGPALSCGNTVVVKPAEQTPLTALHVASLIKEAGFPFPGV
 VNIVPGYGPTAGAAISSHMDIDKVAFTGSTEVGKLIKEAAGKSNLKRVTLELGGKSPC
 IVLADADLDNAVEFAHHGVFYHQGCCIAASRIFVEESIYDEFVRRSVERAKKYILGN
 PLTPGVTQGPQIDKEQYDKILDLIESGKKEGAKLECGGGPWGNKGYFVQPTVFSNVTD
 EMRIAKEEIFGPVQQIMKFKSLDDVIKRANNTFYGLSAGVFTKDIDKAITISSALQAG
 TVWVNCYGVVSAQCPFGGFKMSGNGRELGEYGFHEYTEVKTVTVKISQKNS"

misc feature 82..1488
 /note="aldedh; Region: Aldehyde dehydrogenase family"

variation 362
 /allele="A"
 /allele="G"
 /db_xref="dbSNP:1049981"

variation 1337
 /allele="A"
 /allele="C"
 /db_xref="dbSNP:1803054"

variation 1397
 /allele="A"
 /allele="T"
 /db_xref="dbSNP:1063447"

BASE COUNT 441 a 293 c 391 g 381 t
ORIGIN

1 atgtcatcct caggcacgcc agacttacct gtcctactca ccgatttgaa gattcaatat
 61 actaagatct tcataaacia tgaatggcat gattcagtga gtggcaagaa atttcctgtc
 121 tttaatcctg caactgagga ggagctctgc caggtagaag aaggagataa ggaggatgtt
 181 gacaaggcag tgaaggccgc aagacaggct tttcagattg gatctccgtg gcgtactatg
 241 gatgcttccg agagggggcg actattatac aagttggctg atttaatcga aagagatcgt
 301 ctgctgctgg cgacaatgga gtcaatgaat ggtggaaaac tctattccaa tgcatatctg
 361 agtgatttag caggctgcat caaacattg cgctactgtg cagggtgggc tgacaagatc
 421 cagggccgta caataccaat tgatggaaat ttttttacct atacaagaca tgaacctatt

10/510903

FIG. 11 (3/3)

```
481 ggtgtatgtg gccaaatcat tccttgggaat ttcccgttgg ttatgctcat ttggaagata
541 gggcctgcac tgagctgtgg aaacacagtg gttgtcaaac cagcagagca aactcctctc
601 actgctctcc acgtggcatc ttttaataaaa gaggcagggt ttctctctgg agtagtgaat
661 attgttcctg gttatgggcc tacagcagggt gcagccattt cttctcacat ggatatagac
721 aaagtagcct tcacaggatc aacagagggt ggcaagttga tcaaagaagc tgccgggaaa
781 agcaatctga agagggtgac cctggagctt ggaggaaaaga gcccttgcac tgtgttagct
841 gatgccgact tggacaatgc tgttgaattt gcacaccatg gggatttcta ccaccagggtc
901 cagtgttgta tagccgcac caggattttt gtggaagaat caatttatga tgagtttgtt
961 cgaaggagtg ttgagcgggc taagaagtat atccttggaa atcctctgac cccaggagtc
1021 actcaaggcc ctcagattga caaggaacaa tatgataaaa tacttgacct cattgagagt
1081 gggaagaaaag aaggggccaa actggaatgt ggaggaggcc cgtgggggaa taaaggctac
1141 tttgtccagc ccacagtgtt ctctaattgt acagatgaga tgcgcattgc caaaggaggag
1201 atttttggac cagtgcagca aatcatgaag tttaaatctt tagatgacgt gatcaaaaga
1261 gcaaacaata ctttctatgg cttatcagca ggagtgttta ccaaagacat tgataaagcc
1321 ataacaatct cctctgctct gcaggcagga acagtgtggg tgaattgcta tggcgtggta
1381 agtgcccagt gcccctttgg cggattcaaa atgtctggaa atggaagaga actgggagag
1441 tacggtttcc atgaatatat agagggtcaaa acagtcacag tgaaaatctc tcagaagaac
1501 tcataa
```

//

Revised: October 24, 2001.

Disclaimer | Write to the Help Desk
NCBI | NLM | NIH

10/510903



FIG. 12 (1/2)

Nucleotide

PubMed Nucleotide

Protein

Genome

Structure

PopSet

Taxonomy

OMIM

B

Search Nucleotide

Limits

Preview/Index

History

Clipboard

Details

Display default

Save Text

Add to Clipboard

1: XM_037768. Homo sapiens simi...[gi:14750404]

Related Sequences, Protein, Taxonomy, LinkOut

LOCUS XM_037768 2282 bp mRNA linear PRI 07-FEB-2002
 DEFINITION Homo sapiens similar to pyruvate kinase, muscle (H. sapiens)
 (LOC145710), mRNA.

ACCESSION XM_037768
 VERSION XM_037768.1 GI:14750404

PYRUVATE KINASE

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2282)

AUTHORS NCBI Annotation Project.

TITLE Direct Submission

JOURNAL Submitted (06-FEB-2002) National Center for Biotechnology

Information, NIH, Bethesda, MD 20894, USA

COMMENT GENOME ANNOTATION REFSEQ: This model reference sequence was
 predicted from NCBI contig NT_010235 by automated computational
 analysis using gene prediction method: BLAST. -Also see:-
 Documentation of NCBI's Annotation Process- Evidence Viewer :
 alignments supporting this model.

FEATURES

Location/Qualifiers

source

1..2282

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="15"

gene

1..2282

/gene="LOC145710"

/note="Located on Accession NT_010235"

/db_xref="InterimID:145710"

CDS

109..1704

/gene="LOC145710"

/note="Located on Accession NT_010235"

/codon_start=1

/product="similar to pyruvate kinase, muscle (H. sapiens)"

/protein_id="XP_037768.1"

/db_xref="GI:14750405"

/translation="MSKPHSEAGTAFIQQTQQLHAAMADTFLEHMCRLDIDSPPITARN

TGIICTIGPASRSVETLKEMIKSGMVARLNFSHGTHEYHAETIKNVRTATESFASDP

ILYRPVAVALDTKGPEIRTGLIKGSGTAEVELKKGATLKITLDNAYMEKCDENILWLD

YKNICKVVEVGSKIYVDDGLISLQVKQKGADFLVTEVENGGSLSKKGVNLPGAAVDL

PAVSEKDIQDLKFGVEQDVMVFASFIRKASDVHEVRKVLGEKGKNIKIISKIENHEG

VRRFDEILEASDGIMVARGDLGIEIPAEEKVFLAQKMMIGRCNRAGKPVICATQMLESM

IKKPRPTRAEGSDVANAVLDGADCIMLSGETAKGDYPLEAVRMQHLIAREAEAAIYHL

QLFEELRLAPITSDPTEATAYGAVEASFKCCSGAIIVLTKSGRSAHQVARYRPRAPI

IAVTRNPQTARQAHLYRGIFPVLCKDPVQEAWAEDVDLRVNFAMNVGKARGFFKKGDV

VIVLTGWRPGSGFTNTMRVVPVP"

misc feature

223..1293

FIG. 12 (2/2)

10/510903

variation /note="PK; Region: Pyruvate kinase, barrel domain"
546
/allele="C"
/allele="T"
/db_xref="dbSNP:10514"
misc_feature 1333..1695
/note="PK_C; Region: Pyruvate kinase, alpha/beta domain"
variation 2168
/allele="C"
/allele="T"
/db_xref="dbSNP:1062430"

BASE COUNT 499 a 646 c 654 g 483 t
ORIGIN

```

1 ggctgaggca gtggctcctt gcacagcagc tgcacgcgcc gtggctccgg atctcttcgt
61 ctttgcagcg tagcccgagt cggtcagcgc cggaggacct cagcagccat gtcgaagccc
121 catagtgaag ccgggactgc cttcattcag acccagcagc tgcacgcagc catggctgac
181 acattccttg agcacatgtg ccgcctggac attgattcac caccatcac agcccggaac
241 actggcatca tctgtaccat tggcccagct tcccgatcag tggagacgtt gaaggagatg
301 attaagtctg gaatgaatgt ggctcgtctg aacttctctc atggaactca tgagtaccat
361 gcggagacca tcaagaatgt gcgcacagcc acggaaaagct ttgcttctga cccatcctc
421 taccggcccc ttgctgtggc tctagacact aaaggacctg agatccgaaac tgggtcctc
481 aagggcagcg gcactgcaga ggtggagctg aagaaggag ccactctcaa aatcacgtg
541 gataacgcct acatggaaaa gtgtgacgag aacatcctgt ggctggacta caagaacatc
601 tgcaagggtg tggaagtggg cagcaagatc tacgtggatg atgggcttat ttctctccag
661 gtgaagcaga aaggtgccga cttcctgggtg acggagggtg aaaatgggtg ctccttgggc
721 agcaagaagg gtgtgaacct tcctggggct gctgtggact tgcctgctgt gtcggagaag
781 gacatccagg atctgaagtt tggggtcgag caggatgttg atatgggtgt tgcgtcattc
841 atccgcaagg catctgatgt ccatgaagtt aggaaggctc tgggagagaa gggaaagaac
901 atcaagatta tcagcaaaat cgagaatcat gagggggttc ggaggtttga tgaaatcctg
961 gaggccagtg atgggatcat ggtggctcgt ggtgatctag gcattgagat tcctgcagag
1021 aaggtcttcc ttgctcagaa gatgatgatt ggacgggtgca accgagctgg gaagcctgtc
1081 atctgtgcta ctcagatgct ggagagcatg atcaagaagc cccgccccac tcgggctgaa
1141 ggcaagtgat tggccaatgc agtcctggat ggagccgact gcacatgct gtctggagaa
1201 acagccaaag gggactatcc tctggaggct gtgcgcagtc agcacctgat tgcccgtgag
1261 gcagaggctg ccacttacca cttgcaatta tttgaggaac tccgccgctt ggcgccatt
1321 accagcgacc ccacagaagc caccgccgtg ggtgccgttg aggcctcctt caagtgtgc
1381 agtggggcca taatcgtcct caccaagtct ggcaggctg ctcaccaggt ggccagatac
1441 cgccacagtg ccccatcat tgcgttgacc cggaatcccc agacagctcg tcaggccac
1501 ctgtaccgtg gcactctccc tgtgctgtgc aaggaccagc tccaggaggc ctgggctgag
1561 gacgtggacc tccgggtgaa ctttgccatg aatgttgga agggccgagg cttcttcaag
1621 aagggagatg tggtcattgt gctgaccgga tggcgccctg gctccggctt caccaacacc
1681 atgctgtgtg ttctctgtgc gtgatggacc ccagagcccc tctccagcc cctgtcccac
1741 ccccttcccc cagcccatcc attaggccag caacgcttgt agaactcact ctgggctgta
1801 acgtggcact ggtagggttg gacaccaggg aagaagatca acgcctcact gaaacatggc
1861 tgtgtttgca gcctgctcta gtgggacagc ccagagcctg gctgcccatt atgtggcccc
1921 acccaatcaa gggaagaagg aggaatgctg gactggagge ccttggagcc agatggcaag
1981 aggggtgacag cttcctttcc tgtgtgtact ctgtccagtt ccttttagaaa aaatggatgc
2041 ccagaggact cccaaccctg gcttggggtc aagaaacagc cagcaagagt taggggcctt
2101 agggcactgg gctgtgtgtc cattgaagcc gactctggcc ctggccctta cttgtcttc
2161 tagctctcta ggcctctcca gtttgcacct gtccccaccc tccactcagc tgcctgcag
2221 caaacactcc accctccacc ttccattttc cccactact gcagcacctc caggcctgtt
2281 gc

```

//

Revised: October 24, 2001.

10/510903



FIG. 13 (1/3)



Nucleotide

☐ PubMed ☒ Nucleotide ☐ Protein ☐ Genome ☐ Structure ☐ PopSet ☐ Taxonomy ☐ OMIM ☐ B

Search for

1: XM_049337. Homo sapiens gluc...[gi:14768486] [Related Sequences](#), [Protein](#), [Taxonomy](#), [LinkOut](#)

LOCUS XM_049337 2631 bp mRNA linear PRI 07-FEB-2002
 DEFINITION Homo sapiens glucose-6-phosphate dehydrogenase (G6PD), mRNA.
 ACCESSION XM_049337
 VERSION XM_049337.1 GI:14768486

G6PD

KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2631)
 AUTHORS NCBI Annotation Project.
 TITLE Direct Submission
 JOURNAL Submitted (06-FEB-2002) National Center for Biotechnology
 Information, NIH, Bethesda, MD 20894, USA

COMMENT GENOME ANNOTATION REFSEQ: This model reference sequence was
 predicted from NCBI contig NT_025965 by automated computational
 analysis using gene prediction method: BLAST. ~Also see:~
 Documentation of NCBI's Annotation Process- Evidence Viewer -
 alignments supporting this model.

FEATURES Location/Qualifiers
 source 1..2631
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="X"
 gene 1..2631
 /gene="G6PD"
 /note="G6PD1; Located on Accession NT_025965"
 /db_xref="LocusID:2539"
 /db_xref="MIM:305900"
 CDS 475..2022
 /gene="G6PD"
 /note="Located on Accession NT_025965"
 /codon_start=1
 /product="glucose-6-phosphate dehydrogenase"
 /protein_id="XP_049337.1"
 /db_xref="GI:14768487"
 /translation="MAEQVALSRTQVCGILREELFQGDAFHQSDTHIFIIMGASGDLA
 KKKIYPTIWWLFRDGLLPENTFIMGYARSRLTVADIRKQSEPFFKATPEEKLKLEDF
 ARNSYVAGQYDDAASYQRLNSHMDALHLGSQANRLFYALPPTVYEAVTKNIHESCMS
 QIGWNRIIVEKPFGRDLQSSDRLSNHISLRFREDQIYRIDHYLGKEMVQNLMLVLRFAN
 RIFGPIWNRDNIACVILTFKEPFGTEGRGGYFDEFGIIRDVMQNHLLOMLCLVAMEKP
 ASTNSDDVRDEKVKVLCISEVQANNVVLGGYVGNPDGEGEATKGYLDDPTVPRGSTT
 ATFAAVVLYVENERWDGVPFILRCGKALNERKAEVRLQFHDVAGDIFHQQCKRNLVI
 RVQPNEAVYTKMMTKKPGMFFNPEESELDTYGNRYKNVKLPDAYERLILDVFCGSQM
 HFVRSDELREAWRIFTPLHQLIELEKPKPIPIYGSRGPTADELMKRVGFQYEGTYK
 WVNPHKL"
 variation 507

FIG. 13 (2/3)

10/510903

```

/allele="C"
/allele="G"
/db_xref="dbSNP:1050827"
misc_feature 553..1104
              /note="G6PD; Region: Glucose-6-phosphate dehydrogenase,
              NAD binding domain"
variation 676
          /allele="A"
          /allele="G"
          /db_xref="dbSNP:1050828"
variation 850
          /allele="A"
          /allele="G"
          /db_xref="dbSNP:1050829"
misc_feature 1108..1992
              /note="G6PD_C; Region: Glucose-6-phosphate dehydrogenase,
              C-terminal domain"
variation 2379
          /allele="A"
          /allele="G"
          /db_xref="dbSNP:1050757"
variation 2392
          /allele="A"
          /allele="G"
          /db_xref="dbSNP:1063529"
variation 2490
          /allele="A"
          /allele="G"
          /db_xref="dbSNP:1050830"
variation 2553
          /allele="C"
          /allele="T"
          /db_xref="dbSNP:1050773"
variation 2555
          /allele="C"
          /allele="T"
          /db_xref="dbSNP:1050774"
variation 2593
          /allele="C"
          /allele="T"
          /db_xref="dbSNP:1050831"

```

```

BASE COUNT      527 a      884 c      797 g      423 t
ORIGIN

```

```

1 agggacagcc cagaggaggc gtggccacgc tgccggcgga agtggagccc tccgcgagcg
61 cgcgaggccg ccggggcagg cggggaaacc ggacagtagg ggcggggccc ggccggcgat
121 ggggatgcgg gagcactacg cggagctgca cccgtgcccg ccggaattgg ggatgcagag
181 cagcggcagc gggatatggc ggcagccggc gggccggcct ccagcgcagg tgcccgagag
241 gcaggggctg gcctgggatg cgcgcgcacc tgccctcgcc ccgccccgcc cgcacgaggg
301 gtggtggccg aggccccgcc ccgcacgcct cgcctgaggc ggggtccgctc agcccaggcg
361 cccgcccccg cccccgccga ttaaattggc cggcggggct cagcccccg aaacggtcgt
421 acacttcggg gctgcgagcg cggagggcga cgacgacgaa gcgcagacag cgtcatggca
481 gagcaggtgg ccctgagccg gacccaggtg tgccggatcc tgccgggaaga gcttttccag
541 ggcgatgcct tccatcagtc ggatacacac atattcatca tcatgggtgc atcgggtgac
601 ctggccaaga agaagatcta cccaccatc tggtggctgt tccgggatgg ccttctgccc
661 gaaaacacct tcatcatggg ctatgccgtg tcccgcctca cagtggctga catccgcaaa
721 cagagtgagc ccttcttcaa ggccacccca gaggagaagc tcaagctgga ggacttcttt
781 gcccgaact cctatgtggc tggccagtac gatgatgcag cctcctacca gcgcctcaac
841 agccacatgg atgccctcca cctggggtea caggccaacc gcctcttcta cctggccttg
901 cccccgaccg tctacgagcg cgtcaccaag aacattcacg agtcctgcat gagccagata
961 ggctggaacc gcatcatcgt ggagaagccc ttcgggaggg acctgcagag ctctgaccgg

```

10 / 510903

FIG. 13 (3/3)

```
1021 ctgtccaacc acatctcctc cctgttccgt gaggaccaga tctaccgcat cgaccactac
1081 ctgggcaagg agatggtgca gaacctcatg gtgctgagat ttgccaacag gatcttcggc
1141 cccatctgga accgggacaa catcgctcgc gttatctca ccttcaagga gccctttggc
1201 actgagggtc gcgggggcta ttctgatgaa ttgggatca tccgggacgt gatgcagaac
1261 cacctactgc agatgctgtg tctggtggcc atggagaagc ccgcctccac caactcagat
1321 gacgtccgtg atgagaaggc caaggtgttg aaatgcatct cagaggtgca ggccaacaat
1381 gtggtcctgg gccagtacgt ggggaacccc gatggagagg gcgaggccac caaagggtag
1441 ctggacgacc ccacggtgcc ccgcgggtcc accaccgcca cttttgcagc cgtcgtcctc
1501 tatgtggaga atgagaggtg ggatgggggt cccttcaccc tgcgctgcgg caaggccctg
1561 aacgagcgca aggccgaggt gaggctgcag ttccatgatg tggccggcga catcttccac
1621 cagcagtgca agcgcaacga gctggtgatc cgcgtgcagc ccaacgaggc cgtgtacacc
1681 aagatgatga ccaagaagcc gggcatgttc ttcaaccccg aggagtcgga gctggacctg
1741 acctacggca acagatacaa gaacgtgaag ctccctgacg cctacgagcg cctcactcctg
1801 gacgtcttct gcgggagcca gatgcacttc gtgcgagcg acgagctccg tgaggcctgg
1861 cgtattttca cccactgct gcaccagatt gagctggaga agcccaagcc catccccctat
1921 atttatggca gccgaggccc cacggaggca gacgagctga tgaagagagt gggtttccag
1981 tatgagggca cctacaagtg ggtgaacccc cacaagctct gagecctggg caccacctc
2041 ccccccgcc acggccaccc tccttcccgc cggccgaccc cgagtcggga ggactccggg
2101 accattgacc tcagctgcac attcctggcc ccgggctctg gccaccctgg cccgccccctc
2161 gctgctgcta ctaccgagc ccagctacat tcctcagctg ccaagcactc gagaccatcc
2221 tggccccctc agaccctgcc tgagcccagg agctgagtca cctcctccac tactccagc
2281 ccaacagaag gaaggaggag ggcgcccatt cgtctgtccc agagcttatt ggccactggg
2341 tctcactcct gaggggggcc aggggtggag ggagggacaa gggggaggaa agggcgagc
2401 acccacgtga gagaatctgc ctgtggcctt gcccgccagc ctcagtgcca cttgacattc
2461 cttgtcacca gcaacatctc gagccccctg gatgtccctt gtcccaccaa ctctgcactc
2521 catggccacc ccgtgccacc cgtaggcagc ctctctgcta taagaaaagc agacgcagca
2581 gctgggaccc ctcccaacct caatgccctg ccattaaatc cgcaaacagc c
```

//

Revised: October 24, 2001.

Disclaimer | Write to the Help Desk
NCBI | NLM | NIH

10/510903

FIG. 14 (1/2)



Nucleotide

PubMed Nucleotide Protein Genome Structure PopSet Taxonomy OMIM 8

Search for

Limits Preview/Index History Clipboard Details

1: XM_049047[gi:14759750]

LOCUS XM_049047 1564 bp mRNA linear PRI 16-JUL-2001
DEFINITION Homo sapiens proliferation-associated 2G4, 38kD (PA2G4), mRNA.
ACCESSION XM_049047
VERSION XM_049047.1 GI:14759750
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1564)
AUTHORS NCBI Annotation Project.
TITLE Direct Submission
JOURNAL Submitted (12-JUL-2001) National Center for Biotechnology
Information, NIH, Bethesda, MD 20894, USA

HCDR-3

FEATURES Location/Qualifiers
source 1..1564
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="12"
gene 1..1564
/gene="PA2G4"
/db_xref="LocusID:5036"
/db_xref="MIM:602145"
CDS 120..1304
/gene="PA2G4"
/codon_start=1
/product="proliferation-associated 2G4, 38kD"
/protein_id="XP_049047.1"
/db_xref="GI:14759751"
/translation="MSGEDEQQEQTIAEDLVVTKYKMGDIANRVLRLSLEASSSGVS
VLSLCEKGDAMIMEETGKIFKKEKEMKKGIAFPTSISVNNVCVCHFSPKSDQDYILKE
GDLVKIDLGHVVDGFIANVAHTFVVDVAQGTQVTGRKADVIKAAHLCAEALRLVKPG
NQNTQVTEAWNKVAFSNCPTIEGMLSHQLKQHVIDGEKTI IQNPTDQKKDHEKAEF
EVHEVYAVDVLVSSGEGKAKDAGQRTTIYKRDPKQYGLKMKTSRAFFSEVERRFDAM
PFTLRAFEDKKARMGVVECAKHELLQPFNVLYEKEGEFVAQFKFTVLLMPNGPMRIT
SGPFEPDLYKSEMEVQDAELKALLQSSASRKTQKKKKKASKTAENATSGETLEENEA
GD"

BASE COUNT 455 a 365 c 413 g 331 t
ORIGIN

```
1 ctttcgctcg ccctctctc gaggatcgag gggactctga ccacagcctg tggctgggaa
61 gggagacaga ggcggcggcg gctcagggga aacgaggctg cagtgggtgg agtaggaaga
121 tgtcggggcg ggacgagcaa caggagcaaa ctatcgctga ggacctgggc gtgaccaagt
181 ataagatggg gggcgacatc gccaacaggg tacttcgggc cttgggtggaa gcatctagct
241 cagggtgtgc ggtactgagc ctgtgtgaga aagggtgatgc catgattatg gaagaaacag
301 ggaaaatctt caagaaagaa aaggaaatga agaaaggatg tgcttttccc accagcattt
361 cggtaaataa ctgtgtatgt cacttctccc ctttgaagag cgaccaggat tatattctca
421 aggaagggtga cttggtaaaa attgaccttg ggggccatgt ggatggcttc atcgctaattg
```

10/510903

FIG. 14 (2/2)

```
481 tagctcacac ttttgtggtt gatgtagctc aggggaccca agtaacaggg aggaaagcag
541 atgttattaa ggcagctcac ctttgtgctg aagctgccct acgcctgggc aaacctggaa
601 atcagaacac acaagtgaca gaagcctgga acaaagttgc ccactcattt aactgcacgc
661 caatagaagg tatgctgtca caccagttga agcagcatgt catcgatgga gaaaaaacca
721 ttatccagaa tcccacagac cagcagaaga aggaccatga aaaagctgaa tttgaggtag
781 atgaagtata tgctgtggat gttctcgtca gctcaggaga gggcaaggcc aaggatgcag
841 gacagagAAC cactatttac aaacgagacc cctctaaaca gtatggactg aaaatgaaaa
901 cttcacgtgc cttcttcagt gaggtggaaa ggcgttttga tgccatgccg tttactttaa
961 gagcatttga agatgagaag aaggctcgga tgggtgtggt ggagtgcgcc aaacatgaac
1021 tgctgcaacc atttaatgtt ctctatgaga aggagggtga atttggtgcc cagttaaata
1081 ttacagttct gctcatgccc aatggcccca tgcggataac cagtgggtccc ttcgagcctg
1141 acctctacaa gtctgagatg gaggtccagg atgcagagct aaaggccctc ctccagagtt
1201 ctgcaagtcg aaaaaccag aaaaagaaaa aaaagaaggc ctccaagact gcagagaatg
1261 ccaccagtgg ggaaacatta gaagaaaatg aagctgggga ctgagggtggg tcccatctcc
1321 ccagcttgct gctcctgcct catccccttc ccaccaaacc ccagactctg tgaagtgcag
1381 ttctttctcca cctaggaccg ccagcagagc ggggggatct ccctgcccc accccagttc
1441 cccaaccac tcccttccaa caacaaccag ctccaactga ctctggtctt gggagggtgag
1501 gcttcccaac cacggaagac tactttaaat gaaaaaaga aattgaataa taaaatcagg
1561 agtc
```

//

Revised: October 24, 2001.

Disclaimer | Write to the Help Desk
NCBI | NLM | NIH



FIG. 15 (1/2)

10/510903
Nucleotide

PubMed

Nucleotide

Protein

Genome

Structure

PopSet

Taxonomy

OMIM

8

Search Nucleotide E for

Go

Clear

Limits

Preview/Index

History

Clipboard

Details

Display default [Save] [Print] [Add to Clipboard]

1: XM_052326[gi:14748477]

LOCUS XM_052326 3273 bp mRNA linear PRI 16-JUL-2001
 DEFINITION Homo sapiens DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 21 (DDX21), mRNA.

ACCESSION XM_052326
 VERSION XM_052326.1 GI:14748477

DDX21

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 3273)

AUTHORS NCBI Annotation Project.

TITLE Direct Submission

JOURNAL Submitted (12-JUL-2001) National Center for Biotechnology Information, NIH, Bethesda, MD 20894, USA

FEATURES

source

Location/Qualifiers

1..3273

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="10"

gene

1..3273

/gene="DDX21"

/note="GURDB; RH-II/GU"

/db_xref="LocusID:9188"

CDS

35..1711

/gene="DDX21"

/codon_start=1

/product="DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 21"

/protein_id="XP_052326.1"

/db_xref="GI:14748478"

/translation="MPGKLRS DAGLES DTAMKKGETLRKQTEEEKKKPKSKDKTEEI

AEEEEETVFPKAKQVKKKAEPSEVDMNSPKSKKAKKKEEPSQNDISP KTKSLRKKKEPI

EKKVVSSTKKVTKNEEPSEEEIDAPKPKMKKKEKEMNGETREKSPKLKNGFPHPEPD

CNPSEAAASEESNSEIEQEIPVEQKEGAFSNFPISEETIKLLKGRGVTFLFPIQAKTFH

HVYSGKDLIAQARTGTGKTFSFAIPLIEKLHGELQDRKRGRAPQVLVLAPTRELANQV

SKDFSDITKKLSVACFYGGTPYGGQFERMRNGIDILVGTGPGR IKDHIQNGKLDLTKLK

HVVLDDEVQMLDMGFADQVEEILSVAYKKDSEDNPQTLLFSATCPHWVFNVAKKYMK

TYEQVDLIGKKTQKTAITVEHLAIKCHWTQRAAVIGDVIRVYSGHQGR TIIFCETKKE

AQELSQNSAIKQDAQSLHGDIPQKQREITLKGFRNGSFGVLVATNVAARGLDIPEVDL

VIQSSPPKGCRLVHSSIRADRQSWKDGGVHLLLSAQGRISVSTSGAKSGN"

BASE COUNT 1068 a 603 c 773 g 829 t

ORIGIN

```

1 gaagaccggt cggcctgggc aacctgcgct gaagatgccg ggaaaactcc gtagtgacgc
61 tggtttggaa tcagacaccg caatgaaaaa aggggagaca ctgcgaaaagc aaaccgagga
121 gaaagagaaa aaagagaagc caaaatctga taagactgaa gagatagcag aagaggaaga
181 aactgttttc cccaaagcta aacaagttaa aaagaaagca gaggccttctg aagttgacat
241 gaattctcct aaatccaaaa aggcaaaaaa gaaagaggag ccattctcaa atgacatttc

```

FIG. 15 (2/2)

10/510903

301 tcctaaaacc aaaagtttga gaaagaaaaa ggagcccdatt ga'aa'g'aa'g'-tgggtttcttc
 361 taaaacccaaa aaagtgcacaa aaaatgagga gccttctgag gaagaaatag atgctcctaa
 421 gcccaagaag atgaagaaag aaaaggaaat gaatggagaa actagagaga aaagccccaa
 481 actgaagaat ggatttcctc atcctgaacc ggactgtaac cccagtgaag ctgccagtga
 541 agaaagtaac agtgagatag agcaggaaat acctgtggaa caaaaagaag gcgctttctc
 601 taattttccc atatctgaag aaactattaa acttctcaaa ggccgaggag tgaccttctc
 661 atttcttata caagcaaaga cattccatca tgtttacagc gggaaggact taattgcaca
 721 ggcacggaca ggaactggga agacattctc ctttgccatc cttttgattg agaaacttca
 781 tggggaactg caagacagga agagaggccg tgcccctcag gtactggttc ttgcacctac
 841 aagagagttg gcaaatacaag taagcaaaga cttcagtgac atcacaaaaa agctgtcagt
 901 ggcttggtttt tatgggtggaa ctccctatgg aggtcaattt gaacgcatga ggaatgggat
 961 tgatatcctg gttggaacac caggctcgat caaagaccac atacagaatg gcaaactaga
 1021 tctcacccaaa ctttaagcatg ttgtcctgga tgaagtggac cagatgttgg atatgggatt
 1081 tgctgatcaa gtggaagaga ttttaagtgt ggcatacaag aaagattctg aagacaatcc
 1141 ccaaacattg cttttttctg caacttgccc tcattgggta tttaatgttg ccaagaaata
 1201 catgaaatct acatatgaac aggtggacct gattggtaaa aagactcaga aaacggcaat
 1261 aactgtggag catctggcta ttaagtcca ctggactcag agggcagcag ttattgggga
 1321 tgtcatccga gtatatagt gtcatacagg acgcactatc atcttttctg aaaccaagaa
 1381 agaagcccgag gagctgtccc agaattcagc tataaagcag gatgctcagt ccttgcatgg
 1441 agacattcca cagaagcaaa gggaaatcac cctgaaagggt ttagaaatg gtatgtttgg
 1501 agttttggtg gcaaccaatg ttgctgcacg tgggttagac atccctgagg ttgatttggt
 1561 tatacaaagc tctccaccaa agggatgtag agtccatcat tcactgatcc gggcggacag
 1621 gcagagctgg aaggacgggg gtgtgcatct gcttttatca gcacaaggaa gaatatcagt
 1681 tagtacaagt ggagcaaaaa gcgggaatta agttcaaagc aatagggtgt ccttctgcaa
 1741 cagaaataat aaaagcttcc agcaaaagat ccatcaggct tttggattcc gtgcctcca
 1801 ctgccattag tcaacttcaaa caatcagctg agaagctgat agaggagaag ggagctgtgg
 1861 aagctctggc agcagcactg gcccatattt cagggtgccac gtccgtagac cagcgctcct
 1921 tgatcaactc aaatgtgggt tttgtgacca tgatcttgca gtgctcaatt gaaatgccaa
 1981 atattagtta tgcttggaag gaacttaaag agcagctggg cgaggagatt gattccaaag
 2041 tgaagggaat ggtttttctc aaaggaaagc tgggtggttg ctttgatgta cctaccgcat
 2101 cagtaacaga aatacaggag aaatggcatg attcacgacg ctggcagctc tctgtggcca
 2161 cagagcaacc agaactggaa ggaccacggg aaggatatgg aggcttcagg ggacagcggg
 2221 aaggcagctg aggtctcagg ggacagcggg acggaaacag aagattcaga ggacagcggg
 2281 aaggcagtag aggcccgaga ggacagcgat caggagggtg caacaaaagt aacagatccc
 2341 aaaacaaagg ccagaagcgg agtttcagta aagcatttgg tcaataatta gaaatagaag
 2401 atttatatag caaaaagaga atgatgtttg gcaatataga actgaacatt atttttcatg
 2461 caaagttaaa agcacattgt gcctcctttt gaccacttgc caagtccttg tctctttcag
 2521 acacagacaa gcttcattta aattatttca tctgatcatt atcatttata actttattgt
 2581 tacttcttca tcagtttttc cttttgaaag gtgtatgaat tcattacttt tttattctaa
 2641 tgtattatct gtagattaga agataaaatc aagcatgtat ctgcctatac tttgtgagtt
 2701 cacctgtcct tatactcaaa agtgtccctt aatagtgtcc ttccctgaaa taaataccta
 2761 agggagtgtg acagtctctg gaggaccact ttgagccttt ggaagttaag gtttctcag
 2821 ccacctgccg aacagtttct catgtggtcc tattatttgt ctactgagac ttaatactga
 2881 gcaatgtttt gaaacaagat ttcaaactaa tctgggttgt aatacagttt ataccagtgt
 2941 atgctctaga cttggaagat gtagtatgtt tgatgtggat tacctatact tatgttcgtt
 3001 ttgatacatt tttagcttct cattataagg tgattcatgc tttagtgaat tcttcataga
 3061 tagtatatat aaaagtacat tttaatagaa agccagggtt ttaagggaatt tcacatgtat
 3121 aagggtggct catagcttta tttgtaagta ggctggataa atggtgctta aatggtaatg
 3181 tactccactt ctccctattg gaagattaac attatttacc aagaaggact taaggagta
 3241 gggggcgcag attagcattg ctcaagagta tgt

//

Revised: October 24, 2001.

Disclaimer | Write to the Help Desk
 NCBI | NLM | NIH

10/510903



PubMed Nucleotide

Protein

Genome

Structure

PopSet

Taxonomy

OMIM

B

Search Nucleotide for

Limits

Preview/Index

History

Clipboard

Details

 default

1: XM_030607[gi:14786409]

LOCUS XM_030607 2005 bp mRNA linear PRI 16-OCT-2001
 DEFINITION Homo sapiens serine/threonine kinase 15 (STK15), mRNA.

ACCESSION XM_030607

VERSION XM_030607.1 GI:14786409

KEYWORDS

ARK2

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2005)

AUTHORS NCBI Annotation Project.

TITLE Direct Submission

JOURNAL Submitted (11-OCT-2001) National Center for Biotechnology
 Information, NIH, Bethesda, MD 20894, USA

FEATURES

Location/Qualifiers

source

1..2005
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="20"

gene

1..2005
 /gene="STK15"
 /note="BTAK; Located on Accession NT_011362"
 /db_xref="LocusID:8465"
 /db_xref="MIM:603072"

CDS

26..1237
 /gene="STK15"
 /note="Located on Accession NT_011362"
 /codon_start=1
 /product="hypothetical protein XP_030607"
 /protein_id="XP_030607.1"
 /db_xref="GI:14786410"
 /translation="MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQA
 QRVLCPSNSSQRIPLQAQKLVSSHQKPVQKQKQLQATSVPHVSRPLNNTQKSKQPL
 PSAPENNPHEELASKQKNEESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILA
 LKVLFKAQLEKAGVEHQLRREVEIQSHLRHPNLRLYGYFHDATRVYLILEYAPLGTV
 YRELQKLSKFDEQRTATYITELANALSYCHSKRVIHRDIKPENLLGSAGELKIADFG
 WSVHAPSSRRITLCGTLDYLPPEMIEGRMHDEKVDLWSLGVLCYEFVVGKPPFEANTY
 QETYKRISRVEFTFPDFVTEGARDLISRLKHNPSQRPMLREVLEHPWITANSSKPSN
 CQNKESASKQS"

misc feature

422..1174
 /note="S_TKc; Region: Serine/Threonine protein kinases,
 catalytic domain"

misc feature

422..1174
 /note="pkinase; Region: Protein kinase domain"

misc feature

425..1162
 /note="TyrKc; Region: Tyrosine kinase, catalytic domain"

BASE COUNT 585 a 434 c 456 g 530 t

10/510903

FIG. 16 (2/2)

ORIGIN

```

1 cttgggtcct tgggtcgag gcatcatgga ccgatctaaa gaaaactgca ttccaggacc
61 tgttaaggct acagctccag ttggaggtcc aaaacgtgtt ctctgactc agcaatttcc
121 ttgtcagaat ccattacctg taaatagtgg ccaggctcag cgggtcttgt gtccttcaaa
181 ttcttcccag cgcattcctt tgcaagcaca aaagcttgct tccagtcaca agccggttca
241 gaatcagaag cagaagcaat tgcaggcaac cagtgtacct catcctgtct ccaggccact
301 gaataacacc caaaagagca agcagccccct gccatcggca cctgaaaata atcctgagga
361 ggaactggca tcaaaacaga aaaatgaaga atcaaaaaag aggcagtggg ctttgggaaga
421 ctttgaaatt ggtcgccctc tgggtaaaagg aaagtttggg aatgtttatt tggcaagaga
481 aaagcaaagc aagtttattc tggctcttaa agtgttattt aaagctcagc tggagaaagc
541 cggagtggag catcagctca gaagagaagt agaaatacag tcccacctc ggcacacctaa
601 tattcttaga ctgtatgggt atttccatga tgctaccaga gtctacctaa ttctggaata
661 tgcaccactt ggaacagttt atagagaact tcagaaactt tcaaagtttg atgagcagag
721 aactgctact tatataacag aattggcaaa tgccctgtct tactgtcatt cgaagagagt
781 tattcataga gacattaagc cagagaactt acttcttgga tcagctggag agcttaaaat
841 tgcagatttt ggggtggcag tacatgctcc atcttccagg aggaccactc tctgtggcac
901 cctggactac ctgccccctg aaatgattga aggtcggatg catgatgaga aggtggatct
961 ctggagcctt ggagttcttt gctatgaatt tttagttggg aagcctcctt ttgaggcaaa
1021 cacataccaa gagacctaca aaagaatatc acgggttgaa ttcacattcc ctgactttgt
1081 aacagagggg gccaggggacc tcatttcaag actgttgaag cataatccca gccagaggcc
1141 aatgctcaga gaagtacttg aacacccctg gatcacagca aattcatcaa aaccatcaaa
1201 ttgccaaaac aaagaatcag ctagcaaaac gtcttaggaa tcgtgcaggg ggagaaatcc
1261 ttgagccagg gctgccatat aacctgacag gaacatgcta ctgaagttaa ttttaccatt
1321 gactgctgcc ctcaatctag aacgctacac aagaaatatt tgttttactc agcaggtgtg
1381 ccttaacctc cctattcaga aagctccaca tcaataaaca tgacactctg aagtgaaggt
1441 agccacgaga attgtgtctac ttatactggg tcataatctg gaggcaaggt tgcactgcag
1501 ccgccccgtc agcctgtgct aggcattggg tcttcacagg aggcaaatcc agagcctggc
1561 tgtggggaaa gtgaccactc tgccctgacc ccgatcagtt aaggagctgt gcaataacct
1621 tcctagtacc tgagtgaagt tgtaacttat tgggttggcg aagcctggtg aagctgttgg
1681 aatgagtatg tgattctttt taagtatgaa aataaagata tatgtacaga cttgtatttt
1741 ttctctgggt gcattccttt aggaatgctg tgtgtctgtc cggcaccctg gtaggcctga
1801 ttgggtttct agtcctcctt aaccacttat ctcccatatg agagtgtgaa aaataggaac
1861 acgtgctcta cctccattta gggatttgct tgggatacag aagaggccat gtgtctcaga
1921 gctgttaagg gcttattttt ttaaaacatt ggagtcatag catgtgtgta aactttaaat
1981 atgcaaataa ataagtatct atgtc

```

//

Revised: October 24, 2001.

Disclaimer | Write to the Help Desk
NCBI | NLM | NIH

10/510903



FIG. 17 (1/2)

Nucleotide

PubMed

Nucleotide

Protein

Genome

Structure

PopSet

Taxonomy

OMIM

B

Search ☒ Nucleotide ☐ for

Limits

Preview/Index

History

Clipboard

Details

Display

default

☒ Save☐ Text☐ Add to Clipboard

1: BC008442. Homo sapiens,
Sim...[gi:14250074]

Related Sequences, Protein, Taxonomy, UniSTS,
LinkOut

LOCUS BC008442 1584 bp mRNA linear PRI 12-JUL-2001
DEFINITION Homo sapiens, Similar to transmembrane 4 superfamily member 1,
clone MGC:14656 IMAGE:4101110, mRNA, complete:cds.
ACCESSION BC008442
VERSION BC008442.1 GI:14250074
KEYWORDS MGC. TM4 SF1
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1584)

AUTHORS Strausberg, R.

TITLE Direct Submission

JOURNAL Submitted (25-MAY-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: CLONTECH Laboratories, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome

Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www-shgc.stanford.edu>

Contact: (Dickson, Mark) mcd@paxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL Plate: 21 Row: 1 Column: 7

This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.

FEATURES

source

Location/Qualifiers

1..1584

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="MGC:14656 IMAGE:4101110"

/tissue_type="Bone marrow, chronic myelogenous leukemia"

/clone_lib="NIH_MGC_54"

/lab_host="DH10B"

/note="Vector: pDNR-LIB"

CDS

102..710

10/510903

FIG. 17 (2/2)

```

/codon_start=1
/product="Similar to transmembrane 4 superfamily member 1"
/protein_id="AAH08442.1"
/db_xref="GI:14250075"
/translation="MCYGKCARCIGHSLVGLALLCIAANILLYFPNGETKYASENHLS
RFVWFFSGIVGGLLMLLPAFVFIGLEQDDCCGCCGHENCGKRCAMLSSVLAALIGIA
GSGYCVIVAALGLAEGPLCLDSLQWNYTFASTEGQYLLDTSTWSECTEPKHIVEWNV
SLFSILLALGGIEFILCLIQVINGVLGGICGFCCSHQQQYDC"

```

| BASE COUNT | 460 a | 311 c | 337 g | 476 t | | | |
|------------|-------|-------------|-------------|-------------|-------------|-------------|-------------|
| ORIGIN | 1 | gtggtgtttg | ctttctccac | cagaaggcca | cactttcatc | taatttgggg | tatcactgag |
| | 61 | ctgaagacaa | agagaagggg | gagaaaacct | agcagaccac | catgtgctat | gggaagtgtg |
| | 121 | cacgatgcat | cggacattct | ctggtggggc | tcgccctcct | gtgcatcgcg | gctaataattt |
| | 181 | tgctttactt | tcccaatggg | gaaacaaagt | atgcctccga | aaaccacctc | agccgcttcg |
| | 241 | tgtggtttct | ttctggcatc | gtaggagggtg | gcctgctgat | gctcctgcca | gcatttgtct |
| | 301 | tcattgggct | ggaacaggat | gactgctgtg | gctgctgtgg | ccatgaaaac | tgtggcaaac |
| | 361 | gatgtgcgat | gctttcttct | gtattggctg | ctctcattgg | aattgcagga | tctggctact |
| | 421 | gtgtcattgt | ggcagccctt | ggcttagcag | aaggaccact | atgtcttgat | tccctcggcc |
| | 481 | agtggaaacta | cacctttgcc | agcaccgagg | gccagtacct | tctggatacc | tccacatggg |
| | 541 | ccgagtgcac | tgaacccaag | cacattgtgg | aatggaaatgt | atctctgttt | tctatcctct |
| | 601 | tggctccttg | tggaattgaa | ttcatcttgt | gtcttattca | agtaataaat | ggagtgcctg |
| | 661 | gaggcatatg | tggtttttgc | tgtctccacc | aacagcaata | tgactgctaa | aagaaccaac |
| | 721 | ccaggacaga | gccacaatct | tcctctattt | cattgtaatt | tatatatttc | acttgtattc |
| | 781 | atgtgtaaaa | ctttgtatta | gtgtaacata | ctccccacag | tctactttta | caaacgcctg |
| | 841 | taaagactgg | catcttcaca | ggatgtcagt | gtttaaattt | agtaaaacttc | ttttttgttt |
| | 901 | gtttatttgt | ttttgttttt | tttttaggaa | tgaggaaaca | aaccaccctc | tgggggtagt |
| | 961 | ttacagactg | agtgcacagta | ctcagtatat | ctgagataaa | ctctataatg | ttttggataa |
| | 1021 | aaataacatt | ccaatcacta | ttgtatatat | gtgcatgtat | tttttaaatt | aaagatgtct |
| | 1081 | agttgctttt | tataagacca | agaaggagaa | aatccgacaa | cctggaaaga | tttttgtttt |
| | 1141 | cactgcttgt | atgatgtttc | ccattcatac | acctataaat | ctctaacaag | aggccctttg |
| | 1201 | aactgccttg | tgttctgtga | gaaacaaata | tttacttaga | gtggaaggac | tgattgagaa |
| | 1261 | tgttccaatc | caaataaatg | catcacaact | tacaatgctg | ctcattgttg | tgagtactat |
| | 1321 | gagattcaaa | tttttctaac | atatggaaaag | ccttttgtcc | tccaaagatg | agtactaggg |
| | 1381 | atcatgtgtt | taaaaaaaag | aaaggctacg | atgactgggc | aagaagaaag | atgggaaact |
| | 1441 | gaataaagca | gttgatcagc | atcattggaa | catggggacg | agtgcaggca | ggaggaccac |
| | 1501 | gaggaaatac | cctcaaaact | aacttgttta | caacaaaata | aagtattcac | tacgaaaaaa |
| | 1561 | aaaaaaaaaa | aaaaaaaaaa | aaaa | | | |

//

Revised: October 24, 2001.

Disclaimer | Write to the Help Desk
NCBI | NLM | NIH



FIG. 18 (1/2)

10/510903
Nucleotide

PubMed

Nucleotide

Protein

Genome

Structure

PopSet

Taxonomy

OMIM

B

Search for

Limits

Preview/Index

History

Clipboard

Details

Display

1: XM_027538[gi:14768648]

LOCUS XM_027538 1025 bp mRNA linear PRI 16-JUL-2001
 DEFINITION Homo sapiens excision repair cross-complementing rodent repair
 deficiency, complementation group 1 (includes overlapping antisense
 sequence) (ERCC1), mRNA.

ACCESSION XM_027538
 VERSION XM_027538.1 GI:14768648

ERCC1

KEYWORDS

SOURCE

human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 1025)

AUTHORS

NCBI Annotation Project.

TITLE

Direct Submission

JOURNAL

Submitted (12-JUL-2001) National Center for Biotechnology
 Information, NIH, Bethesda, MD 20894, USA

FEATURES

Location/Qualifiers

source

1..1025
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="19"

gene

1..1025
 /gene="ERCC1"
 /note="UV20"
 /db_xref="LocusID:2067"
 /db_xref="MIM:126380"

CDS

63..956
 /gene="ERCC1"
 /codon_start=1
 /product="excision repair cross-complementing rodent
 repair deficiency, complementation group 1 (includes
 overlapping antisense sequence)"
 /protein_id="XP_027538.1"
 /db_xref="GI:14768649"
 /translation="MDPGKDKEGVPQPSGPPARKKFVIPLDEDEVPPGVAKPLFRSTQ
 SLPTVDTSAQAAPQTYAEYAIISQPLEGAGATCPTGSEPLAGETPNQALKPGAKSNSII
 VSPRQRGNPVLKFVRNVPWEFGDVI PDYVLGQSTCALFLSLRYHNLHPDYIHGRLQSL
 GKNFALRVLLVQVDVKDPQQALKELAKMCILADCTLILAWSPEEAGRYLETYKAYEQK
 PADLLMEKLEQDFVSRVTECLTVKSVNKTDSTLLTTFGSLEQLIAASREDLALCPG
 LGPQKARRLFDVLHEPFLKVP"

BASE COUNT 234 a 326 c 289 g 176 t

ORIGIN

```

1 ccaagaccag caggtgaggc ctcgcggcgc tgaaaccgtg aggcccggac cacaggctcc
61 agatggaccc tgggaaggac aaagaggggg tgcccagcc ctcagggccg ccagcaagga
121 agaaatttgt gataccctc gacgaggatg aggtccctcc tggagtggcc aagcccttat
181 tccgatctac acagagcctt cccactgtgg acacctcggc ccaggcggcc cctcagacct
241 acgccgaata tgccatctca cagcctctgg aaggggctgg ggccacgtgc cccacagggt

```

10/510903

FIG. 18 (2/2)

```
301 cagagcccct ggcaggagag acgcccaccc aggcacctgaa acccgggggca aaatccaaca
361 gcatcattgt gagccctcgg cagaggggca atcccgtact gaagttcgtg cgcaatgtgc
421 cctgggaatt tggcgacgta attcccgact atgtgctggg ccagagcacc tgtgccctgt
481 tcctcagcct ccgctaccac aacctgcacc cagactacat ccatgggagg ctgcagagcc
541 tggggaagaa cttcgccctg cgggtcctgc ttgtccagggt ggatgtgaaa gatccccagc
601 aggcctcaa ggagctggct aagatgtgta tcctggccga ctgcacattg atcctcgctt
661 ggagccccga ggaagctggg cggtagctgg agacctaca ggcctatgag cagaaaccag
721 cggacctcct gatggagaag ctagagcagg acttcgtctc ccgggtgact gaatgtctga
781 ccaccgtgaa gtcagtcaac aaaacggaca gtcagacctt cctgaccaca tttggatctc
841 tggaacagct catcgccgca tcaagagaag atctggcctt atgcccaggc ctgggcccctc
901 agaaagcccg gaggtctgtt gatgtcctgc acgagccctt cttgaaagta ccctgatgac
961 ccagctgcc aaggaaaccc ccagtgtaat aataaatcgt cctcccaggc caggctcctg
1021 ctggc
```

//

Revised: October 24, 2001.

Disclaimer | Write to the Help Desk
NCBI | NLM | NIH

10/510903

G2-2F3 // Fanconi Anemia Group A (FANCA)

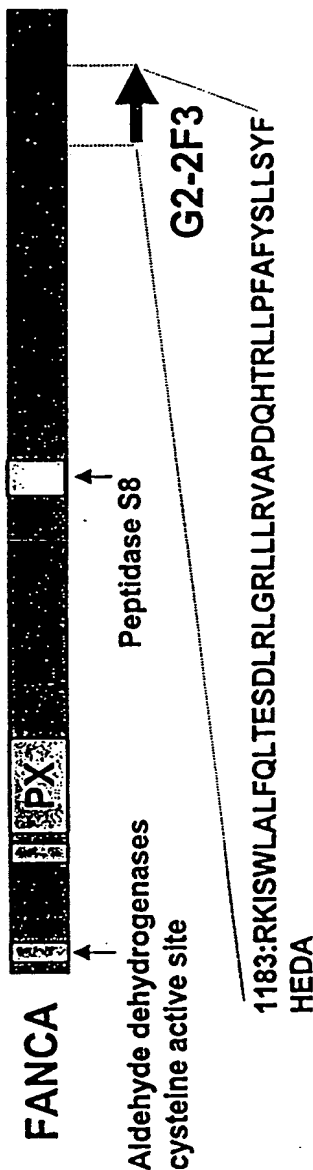
The G2-2F3 sequence is identical to Fanconi Anemia Group A,

FANCA, 1340aa

Orientation : Sense

FKBP-type peptidyl-prolyl
cis-trans isomerase signature 1

FANCA



Pfam HMM search was done at the Washington University web site

Aldehyde dehydrogenases cysteine active site (3-14): It is found in a nuclear protein associated with cell proliferation

FKBP-type peptidyl-prolyl cis-trans isomerase signature 1(159-175): One of two signature patterns for FKBP

PX(189-320): Novel domains in NADPH oxidase subunits, sorting nexins, and PI3-kinases: binding partners of SH3 domains?

PeptidaseS8(660-688): Subtilase family motif

FIG. 19

10/510903

Cell Tracker Analysis of G2-2F3 (The Fragment of FANCA)-Infected A549.tTA Cells

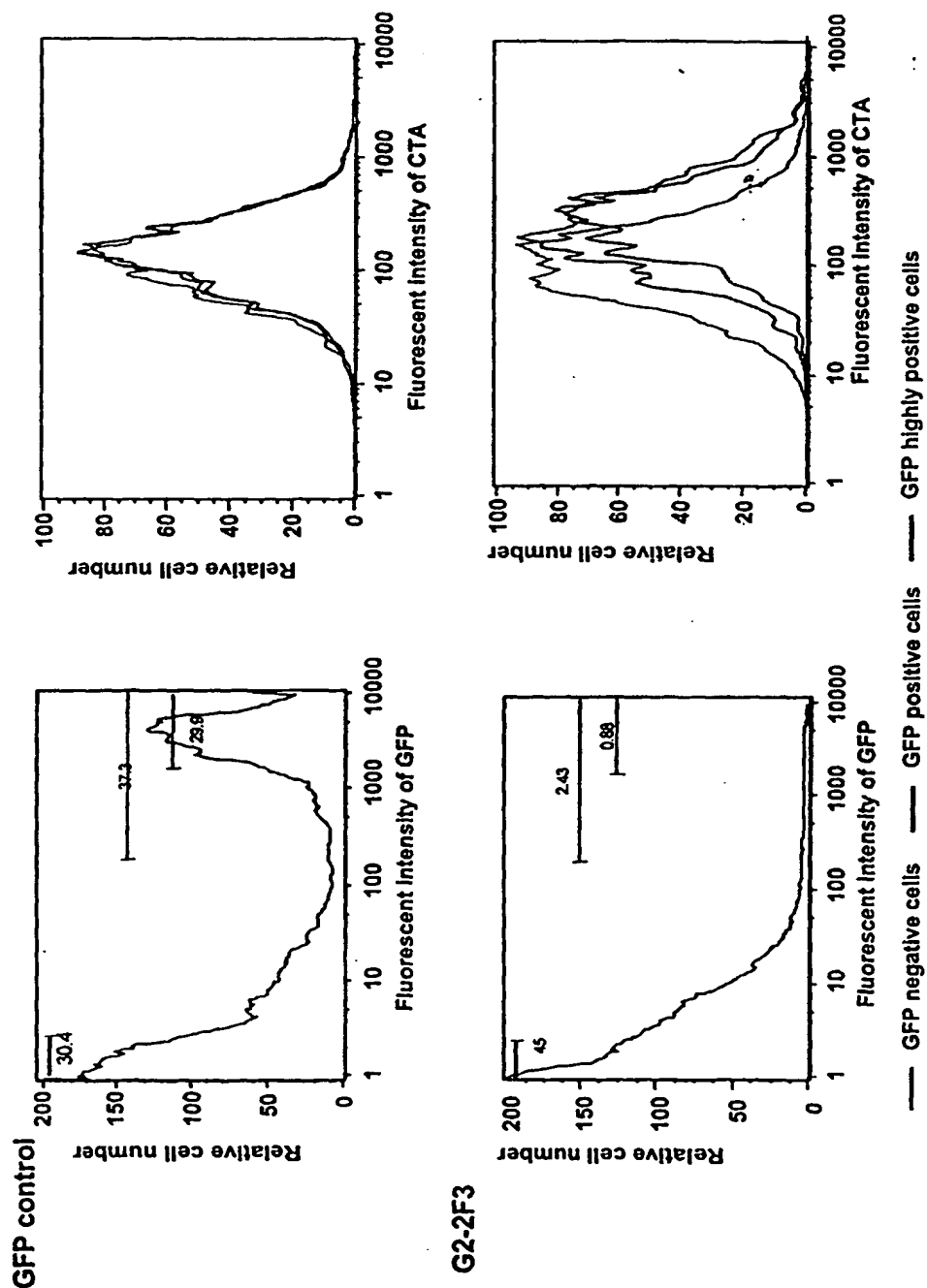


FIG. 20

G3-2H6 // DEAD/H Box Polypeptide 9 (DDX9)

The G3-2H6 sequence is identical to DEAD/H box polypeptide 9 (DDX9), 1279aa
Orientation: Antisense



CLN3[↑] G3-2H6(572bp)
603 bp Insert

C-terminus of GFP

GAGTTCGTGACGCCGCCGGGATCACTTCGGCATGGACGAGCTGTACAAGGAGGAGGCC
GCCAAGGCC
EEFVTAAGITLGMDLYKEEAKA
GGTGGCAGCGTGCTCCAGTGTGCTGGAAAGGCCCACCTCCTCTTCCTGTCCAAAGTA
GCCAGTTCC
GGS GSVLESATSSLSKVASS
ATAGGCCCCCCTACCACCTCCTCGTGAATCCCCAGATCCTCTGTAGCCTCCACTAGGC
CCTCTGTA
IGPPTTSLESPRSSVASTRPSV
GTCTCTCCAGAGTTGCCTAAAGCCACCTCGGAGACTCCTCTATAGCCTCCACCAACA
CCTGCACC
ATATCTGCCCCGAAAGGATTGGCGCTGCCACCA⁷TAGCCTCGCTACCATAGCCTCCACTG
CTATAGCC
ISCPKVGATAIASATAIA
ACCGCATAGCCTCCACCCTGTAAC TAGAACCTCCCTCTATATCGCTTCCATTGTCGTA
ICGGGCC
BBA(50-181-250): Double-stranded RNA binding motif
PEAP(48-669): Amino acids
DEAH box helices
HSC(167-679): Helicase conserved C-terminal domain
GenBank accession numbers are indicated by bold letters.
GATAGCA

FIG. 21

Cell Tracker Analysis of G3-2H6 (The Antisense Fragment of DDX9)-Infected A549.tTA Cells

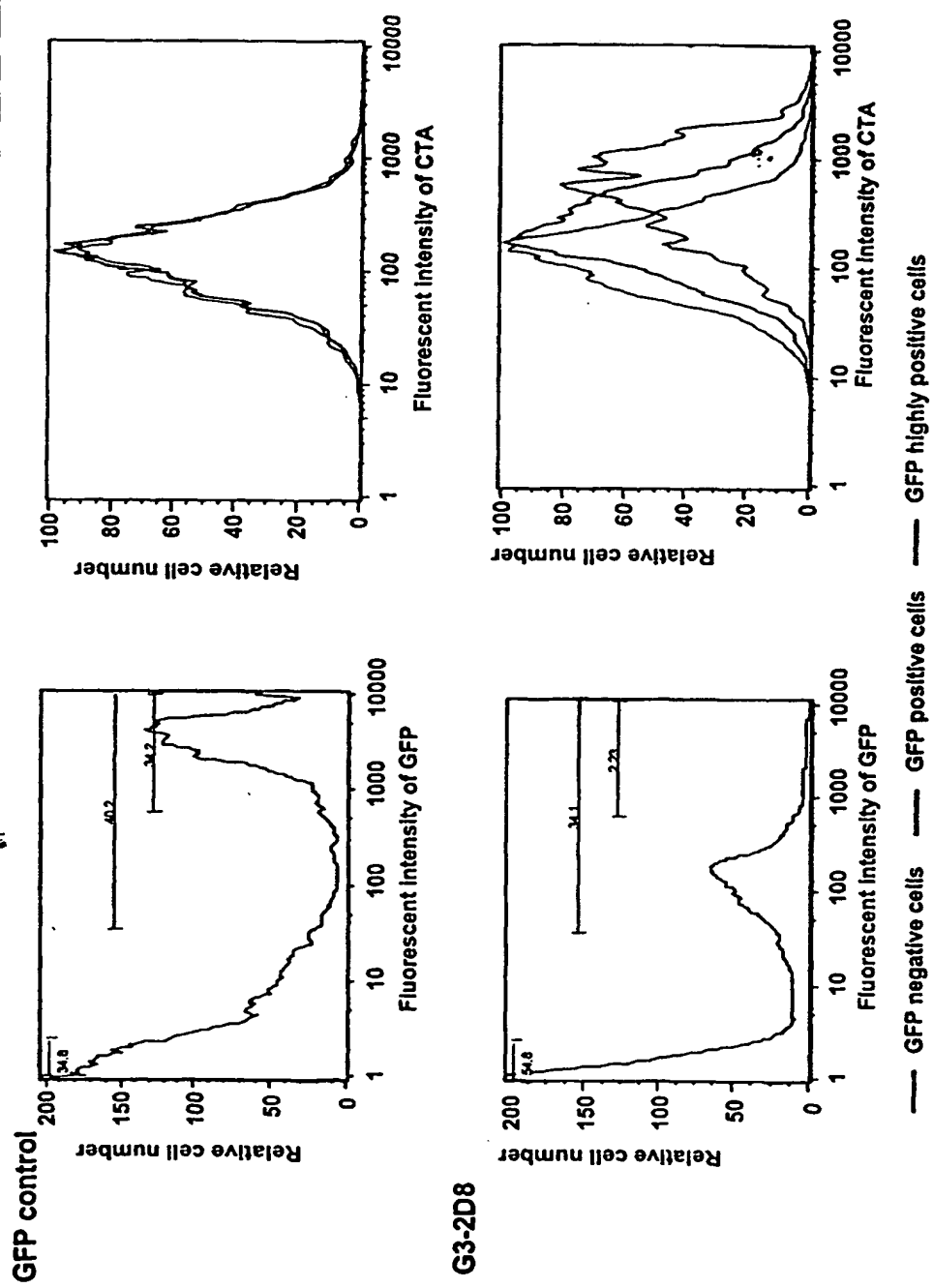


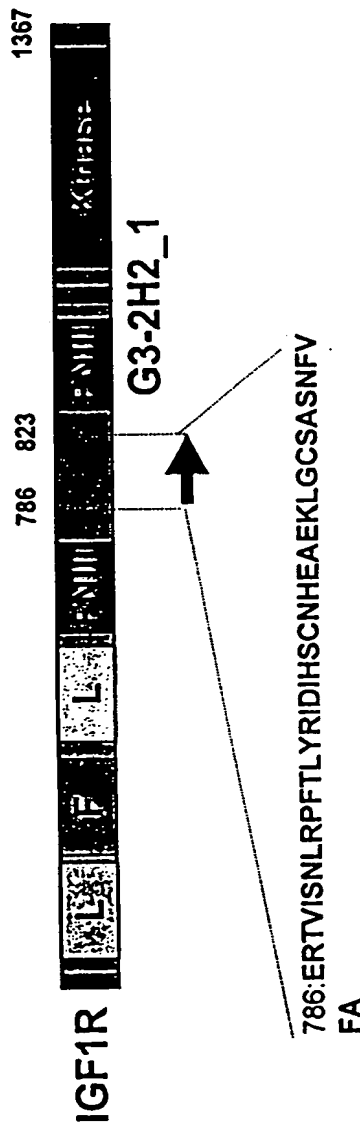
FIG. 22

10/510903

G3-2H2_1/Insulin-like Growth Factor 1 Receptor (IGF1R)

The G3-2H2_1 sequence is identical to Insulin-like growth factor 1 receptor (IGF1R)

Orientation: Sense



Leader sequence (1-30)
 L (51-172, 352-472): Receptor L domain, the L domains from insulin-like growth factor receptors make up the bilobal ligand binding site.
 F (175-333): Furin-like cysteine rich region, which involves receptor aggregation
 FNIII(489-587, 835-917): Fibronectin type III domain, the majority of which are involved in cell surface binding in some manner, or are receptor protein tyrosine kinases, or cytokine receptors.
 Transmembrane (936-958)
 Kinase(999-1266): Protein tyrosine kinase catalytic domain

FIG. 23

Cell Tracker Analysis of G3-2H2_1 (The Fragment of IGF1R)-Infected A549.tTA Cells

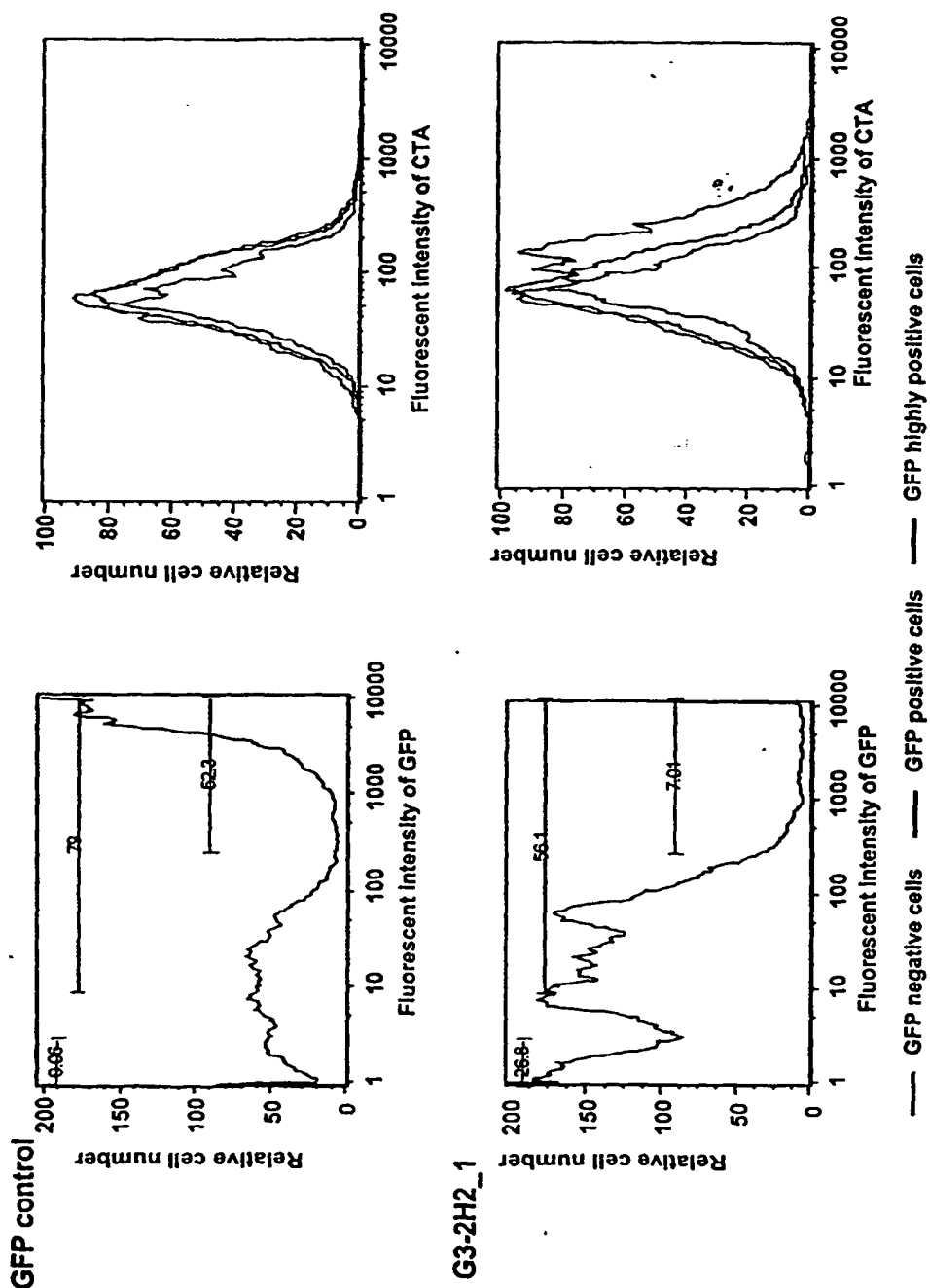


FIG. 24

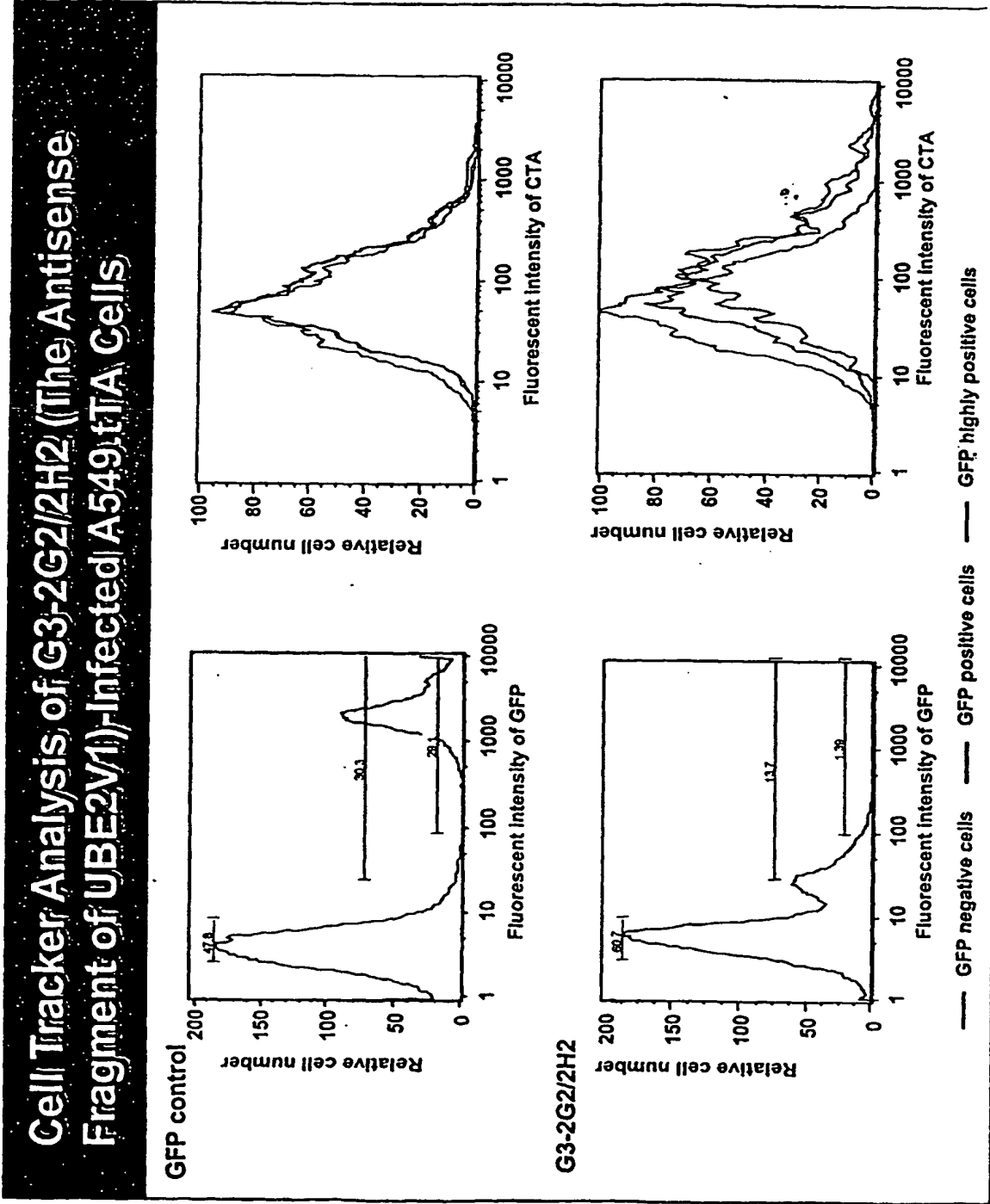


FIG. 26

FIG. 28 (1/2)

10/510903

SEQ ID NO:29

Size: 181

DNA FANCA

CCAGTGTGCTGGAAAGGAGGAAGATATCCTGGCTGGCACTCTTTCAGTTGACAGAGAGTGACCTCAGGCTGGGGC
GGCTCCTCCTCCGTGTGGCCCCGGATCAGCACACCAGGCTGCTGCCTTTTCGCTTTTTACAGTCTTCTCTCCTACT
TCCATGAAGACGCGGCTTTCCAGCACAGTGG

SEQ ID NO:30

Size: 603

DNA DDX9

CCAGTGTGCTGGAAAGCGCCACCTCCTCTTCCCTGTCCAAAGTAGCCAGTTCCATAGGCCCCCTACCACCWCCT
CGCTGGAATCCCCCAGATCCTCTGTAGCCTCCACTAGGCCCTCTGTAGTCTCCTCCAGAGTTGCCTCTAAAGCCA
CCTCGGGAGACTCCTCTATAGCCTCCACCAACACCTGCACCATATCCTGCCCCGAAAGGAGTTGGCGCTGCCACCA
TAGCCTCCGCTACCATAGCCTCCACTGCTATAGCCACCGCATAGCCTCCACCACTGTAAGTAAACCTCCCTTC
TATATCCGCTTCCATTGTCGTATCGGGCCATCTTGGGAGGACGTGGACCATCTCCATGCCGTGTACTGCCAATCA
TAAGGTTGATACCAGCAGCTGAGGGTCTAGAGATCTGACGGATCATGTTTCAGCATACGTTTCATTTACGGGGTCCA
ACTGGCTGATGATAGCAGGTTGTTTGGTTACTTCAACAACCAAAGCCTCCATGGCTGCCCGGAGACCAGTGATAC
AGGCAGCAGCTTCATGAGATATTTGCAGTTTAATCCAGTCATCTACAAGCACAAATCTGCCCACCTTTCCAGCACAG
TGG

SEQ ID NO:31

Size: 145

DNA IGF1R

CCAGTGTGTTGGAAAGGGAGAGAACTGTCATTTCTAACCTTCGGCCTTTCACATTGTACCGCATCGATATCCACA
GCTGCAACCACGAGGCTGAGAAGCTGGGCTGCAGCGCCTCCAACCTTCGTCTTTGCTTTCCAGCACAGTGG

SEQ ID NO:32

Size: 269

DNA UBEV2V1

CCAGTGTGCTGGAAAGGTGCTTCTGGGTATTTAGGTCCACATTCTATTTTAAGGCTGTATATTCGGTTTTTCATAA
ATTGTTCTTGGAGGCCCAATTATCATCCCTGTCCATCTTGTAAGATGTCATGTCTTCGTCATCTTCTAGACCCCA
GCTAACTGTGCCATCTCCTACTCCTTTCTGGCCTTCTTTCGAGATTCTTCCAACAGTCGGAAATTGCGAGGGACTT
TATACATCCCAGCCCCGTGGTGGCTGCCCTTTCCAGCACACTGG

SEQ ID NO:33

Size: 499

DNA aldehyde dehydrogenase

CCAGTGTGCTGGAAAGGAGCAAACCTCTCACTGCTCTCCACGTGGCATCTTTAATAAAAAGAGGCAGGGTTTCC
TCCTGGAGTAGTGAATATTGTTCCCTGGTTATGGGCCTACAGCAGGGGCAGCCATTTCTTCTCACATGGATATAGA
CAAAGTAGCCTTCACAGGATCAACAGAGGTTGGCAAGTTGATCAAAGAAGCTGCCGGGAAAAGCAATCTGAAGAG
GGTGACCCTGGAGCTTGGAGGAAAAGAGCCCTTGCAATTGTGTTAGCTGATGCCGACTTGGACAATGCTGTTGAATT
TGCACACCATGGGGTATTCTACCACCAGGGCCAGTGTGTATAGCCGCATCCAGGATTTTTGTGGAAGAATCAAT
TTATGATGAGTTTTGTTTCAAGGAGTGTTGAGCGGGCTAAGAACGTATATCCTTGGAAACATCCTCTGACCCAG
GAGTCACTCAAAGGCCCTCAGATTGACAAGGACTTTCCAGACACAGTGG

SEQ ID NO:34

Size: 425

DNA pyruvate kinase

400 000 100 000000 000000

FIG. 28 (2/2)

10/510903

CCAGTGTGCTGGAAAGGCTGCCCCACTTCCACCACCTTGCAGATGTTCTTGTAGTCCAGCCACAGGATGTTCTCGT
CACACTTTTCCATGTAGGCGTTATCCAGCGTGATTTTGAGAGTGGCTCCCTTCTTCAGCTCCACCTCTGCAGTGC
CGCTGCCCTTGATGAGCCCAGTTCGGATCTCAGGTCCTTTAGTGTCTAGAGCCACAGCAACGGGCCGGTAGAGGA
TGGGGTCAGAAGCAAAGCTTTCCGTGGCTGTGCGCACATTCTTGATGGTCTCCGCATGGTACTCATGAGTTCCAT
GAGAGAAGTTTCAGACGAGCCACATTCATTCCAGACTTAATCATCTCCTTCAACGTCTCCACTGGATCGGGAAGCT
GGGCCAATGGTACAGATGATGCCAGTGTTCGGGGCTTTCAGCACAGTGG

SEQ ID NO:35

Size:

DNA G6PD

CCAGTGTGCTGGAAACTTTCCAGTTCTCCATGGCCACCANACACAGCATCTGCAGTAGGTGGTTCTGCATCACGT
CCCGGATGATCCCAAATTCATCGAAATAGCCCCCGCGACCCTCAGTGCCAAAGGGCTCCTTGAAGGTGAGGATAA
CGCAGGCGATGTTGTCCCGGTTCCANATGGGGCCGAAGATCCTGTTGGCAAATCTCAGCACCATGAGGTTCTCTT
TCCAGCACAGTGG

10/510903

Dominant Negative Mutants of BAP-1

Point mutants: C91A, H169A- catalytic residues in the protease domain.
(EMBO J. 1997 Jul 1;16(13):3787-96. PMID: 9233788)

CLUSTAL W (1.8) multiple sequence alignment

```

Uch-13      MEGQRWLPLEANPEVTNQFLKQLGLHPNWQFVDVYGMDELLSMVPRPVCVALLFPITE
BAP-1      MNKGWLELESDPGLFTLLVEDFGVGK-VQVEEIIY----DLQSKCQGPVYGFIFLQKWI
           .: ** **: * : . : : : : * : : * : * : * : * : * : * : * :
Uch-13      KYEVFR--TEEEKIKSQGDVTSSVYFMKQTI↓SNACGTIGLIHAIANNKDMHFESGST
BAP-1      ERRSRKVTLVDDTSVIDDDIVNNMFFAHQLIPNSCATHALLSVLLNCSS--VDLGPT
           : . * : : . . . : * : : : * : * : * : * : * : * : * :
Uch-13      LKKFLEESVMSPEERARYLENYDAIRVTHETSAHEGQTEAP-----SIDEKVDLHFI
BAP-1      LSRMKDFTKGFSPESKGYAIGNAPELAKAHNSHARPEPRHLPEKQNGLSAVRTMEAFHFV
           * : : : : : * : : : * : : * : . * : : : * : : * :
Uch-13      ALVHVDGHLIELDGRKPPFPINHGETS-DETLLEDAIEVCKKFMERDPD-----ELRFNAI
BAP-1      SYVPITGRLEFDGLKVYPIDHGPWGGEDEEWTDKARRVIMERIGLATAGEPYHDIRFNL
           : * : * : * : * : * : * : * : * : * : * : * : * : * :

```

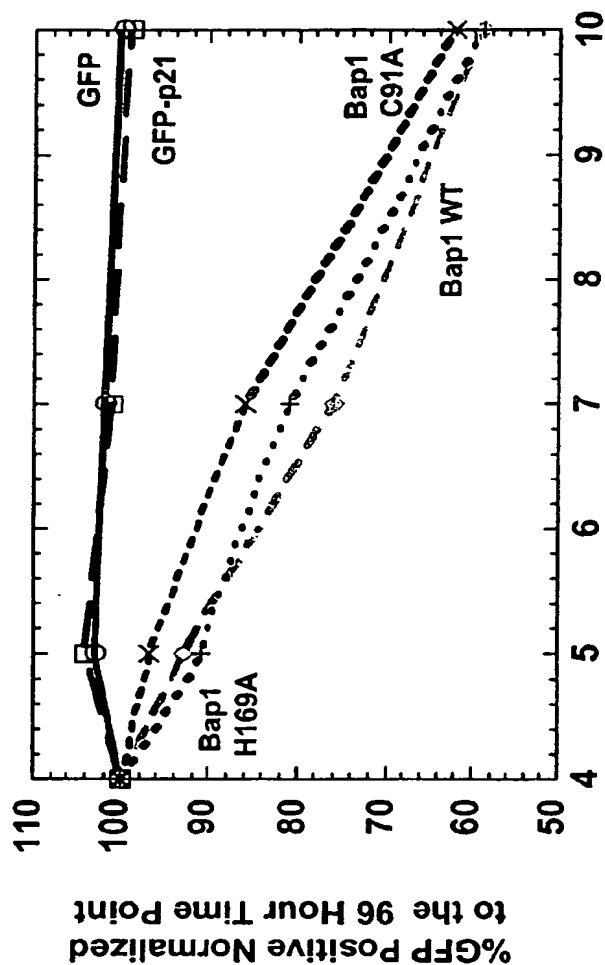
Bold: Catalytic residue

FIG. 29

RICE

10 / 510903

Expression of Bap1 WT and Protease Mutants is Antiproliferative in HeLa Cells



%GFP Positive at 96 Hours

| | | |
|---|------------|------|
| ○ | GFP | 80.8 |
| □ | GFP-p21 | 19.5 |
| ◇ | Bap1 WT | 12.0 |
| × | Bap1 C91A | 14.7 |
| + | Bap1 H169A | 12.5 |

FIG. 30

RIGEL

Expression of Bap1 WT Protein is Antiproliferative in HeLa Cells in the Celltracker Assay

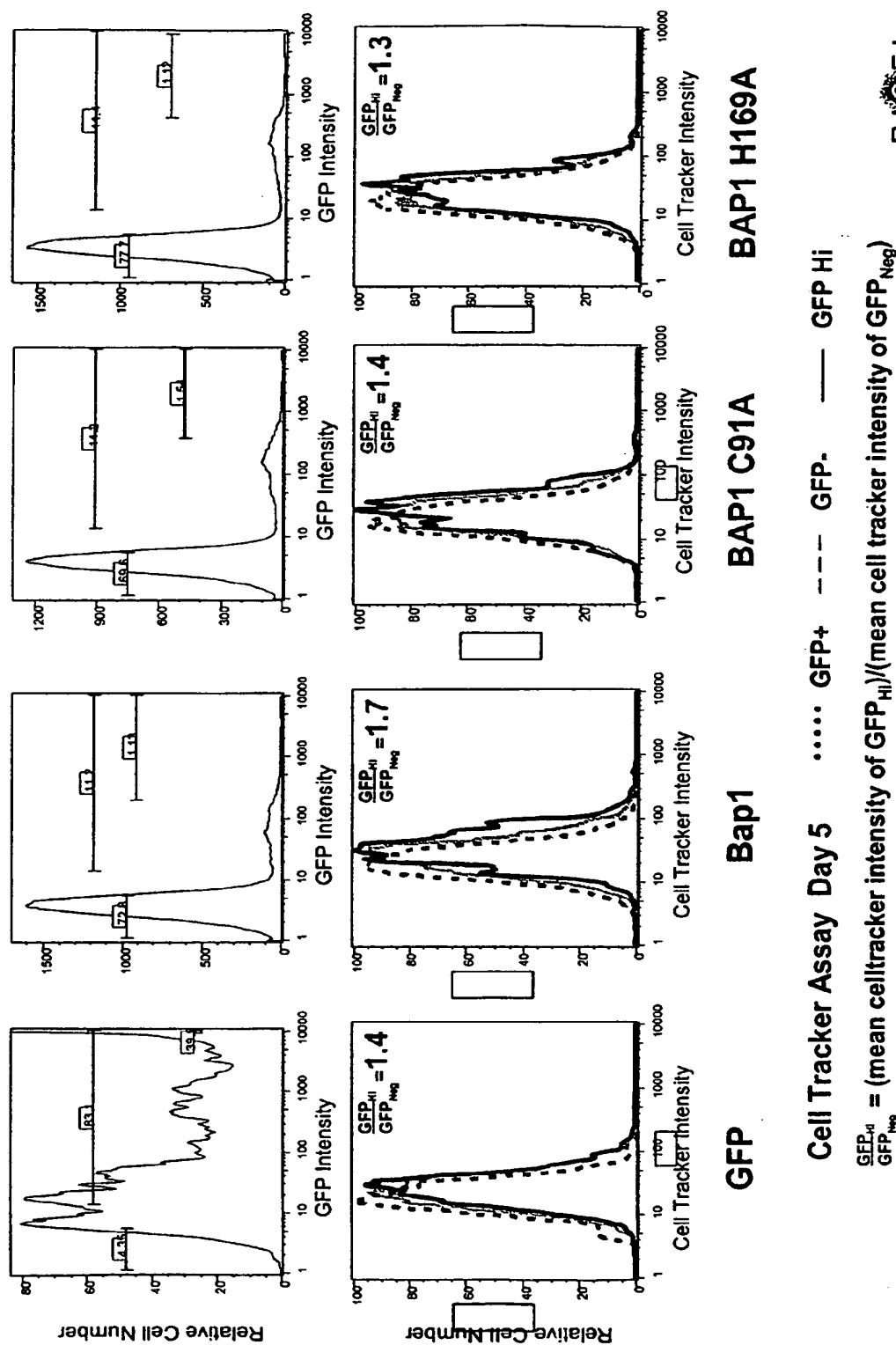


FIG. 31

10/510903

Expression of Bap1 Protease Mutants is Slightly More Antiproliferative Than Expression of Bap1 WT in H1299 Cells

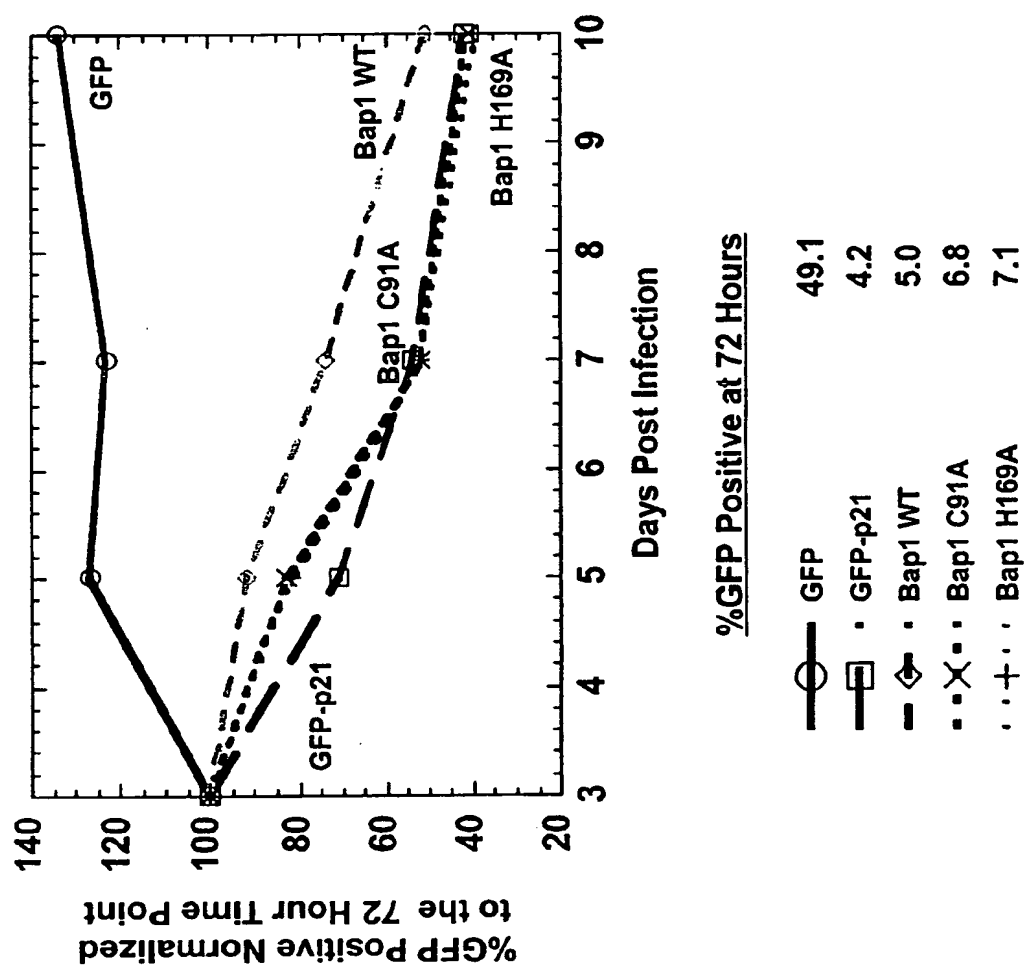
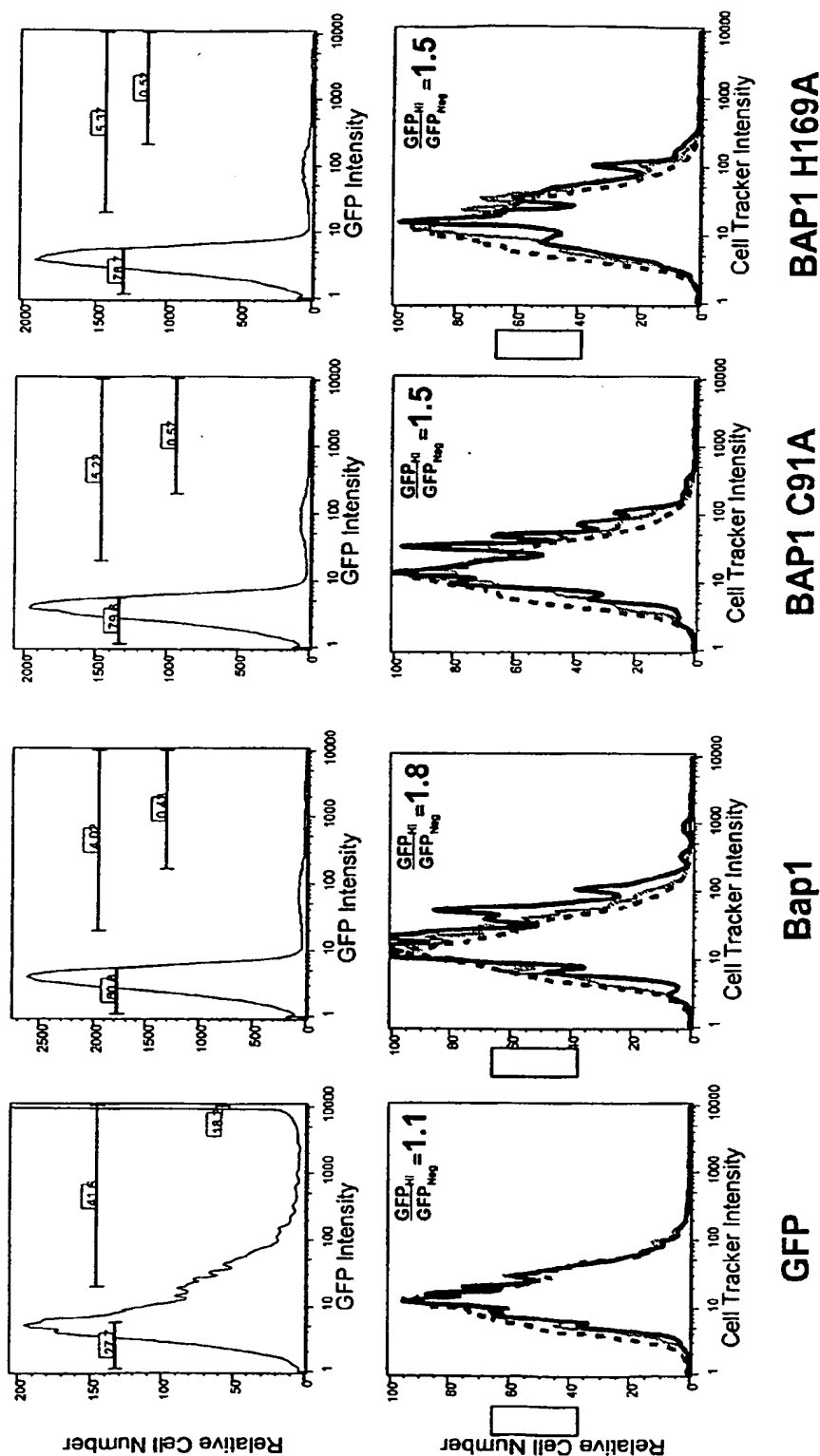


FIG. 32

R 10/510903

Expression of Bap1 WT and Bap1 Protease Mutants is Antiproliferative in H1299 Cells in the Celltracker Assay



Cell Tracker Assay Day 5 GFP+ --- GFP. — GFP Hi

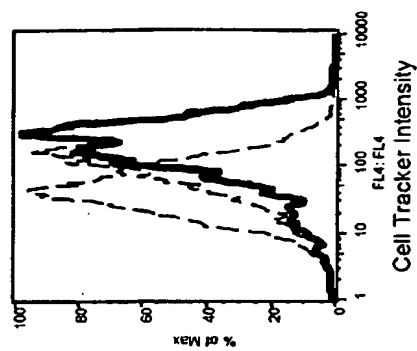
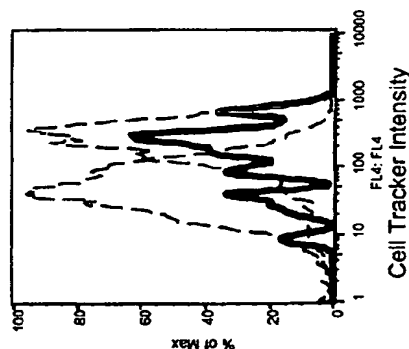
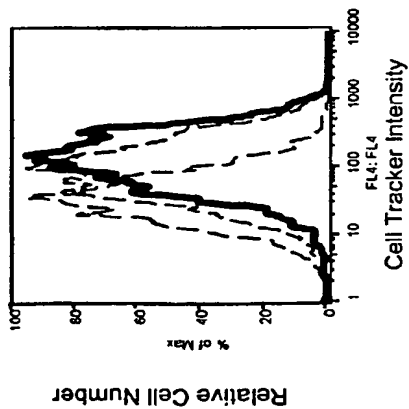
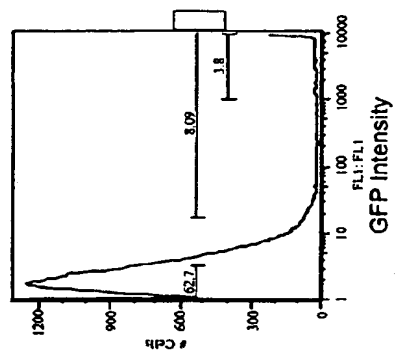
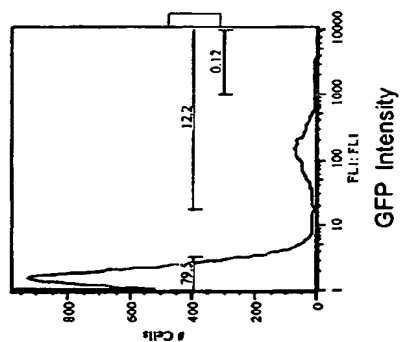
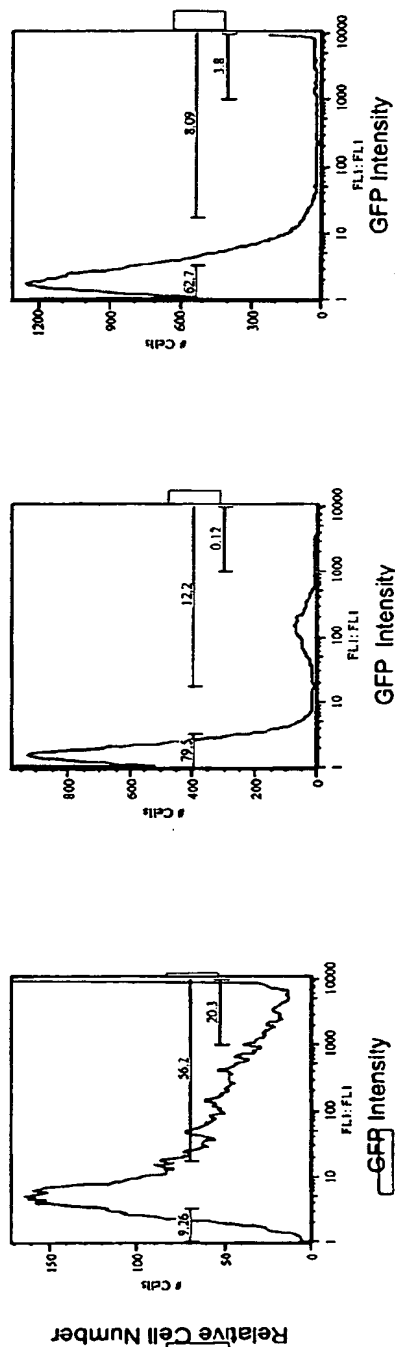
$\frac{GFP_{Hi}}{GFP_{Neg}} = (\text{mean celltracker intensity of } GFP_{Hi}) / (\text{mean cell tracker intensity of } GFP_{Neg})$

FIG. 33

R: G E L

10/510903

The Bap1 Functional Hit G32D8 is Antiproliferative in HMEC Cells



GFP

Gp21

G3-2D8

..... GFP+ --- GFP- — GFP hi

FIG. 34

RIGEL

The Bap1 Functional Hit G3-2D8 is Antiproliferative in PrEC Cells

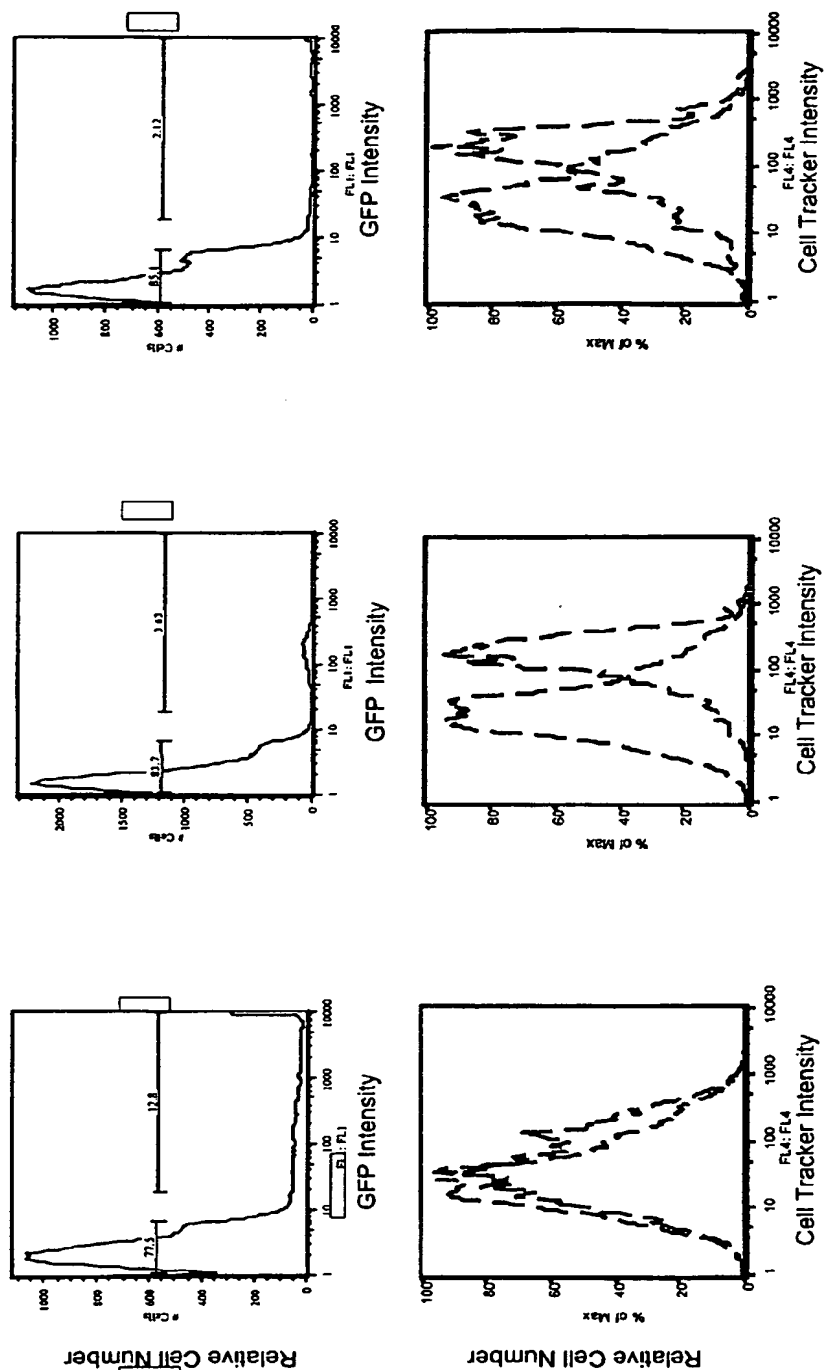
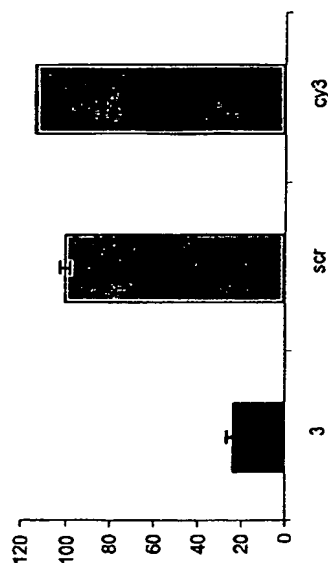


FIG. 35

REL

BAP1 Specific siRNA Has an Antiproliferative Effect on HeLa Cells

BAP1 mRNA levels in HeLa after siRNA treatment (Taqman)



HeLa cell cycle profile after BAP1 siRNA treatment



BrdU incorporation by HeLa treated with BAP1 siRNA

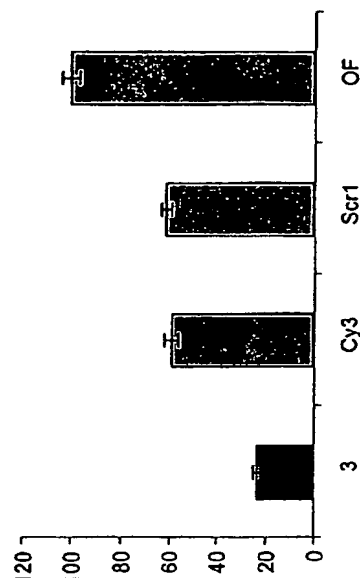
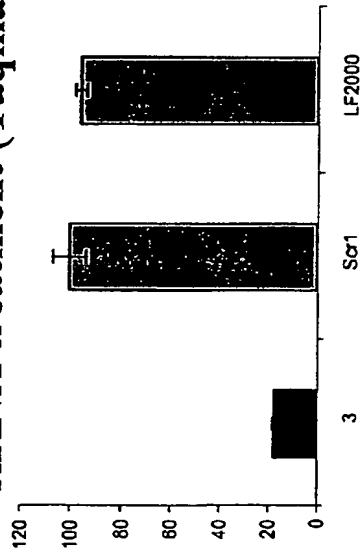


FIG. 36

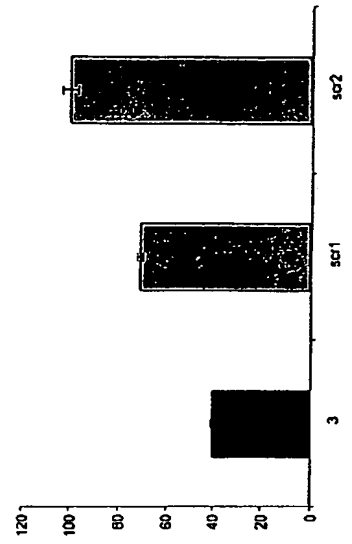
FIG. 36

BAP1 Specific siRNA Induces G1 Arrest in H1299 Cells

BAP1 mRNA levels in H1299 after siRNA treatment (Taqman)



BrdU incorporation by H1299 treated with BAP1 siRNA



H1299 cell cycle profile after BAP1 siRNA treatment

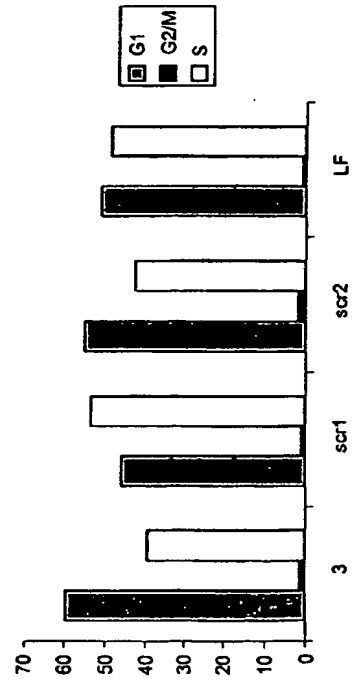
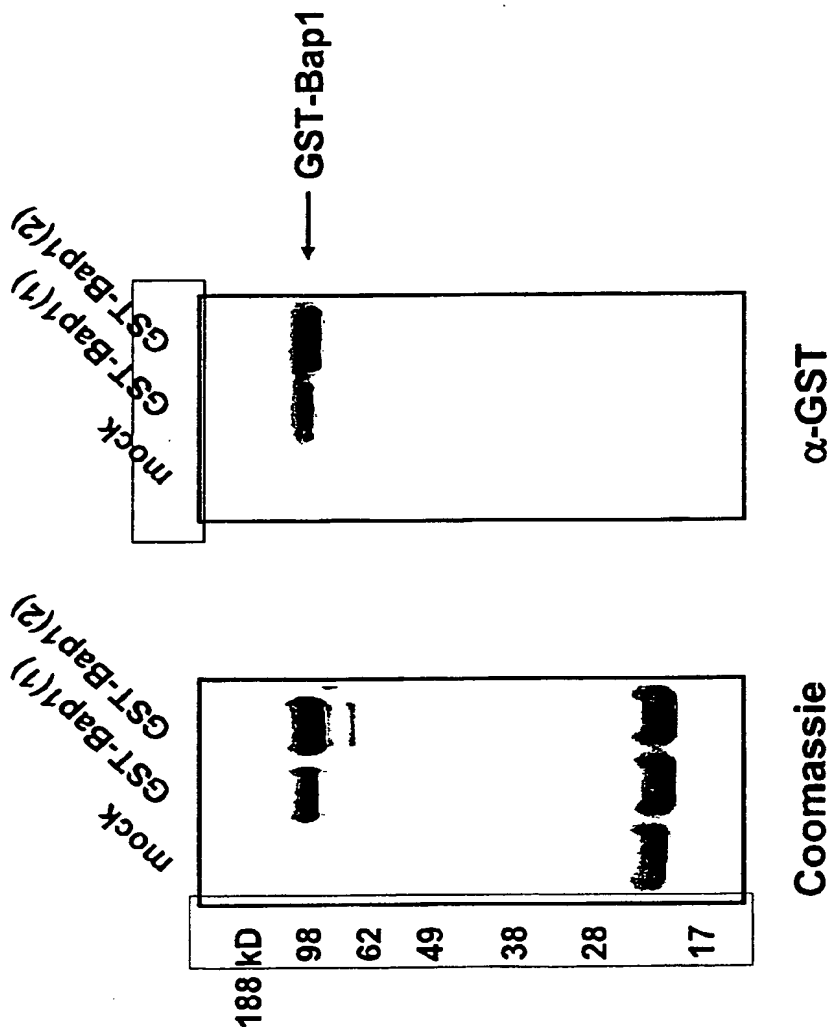


FIG. 37

REL

Soluble GST-Bap1 Protein can be Expressed from SF9 Cells



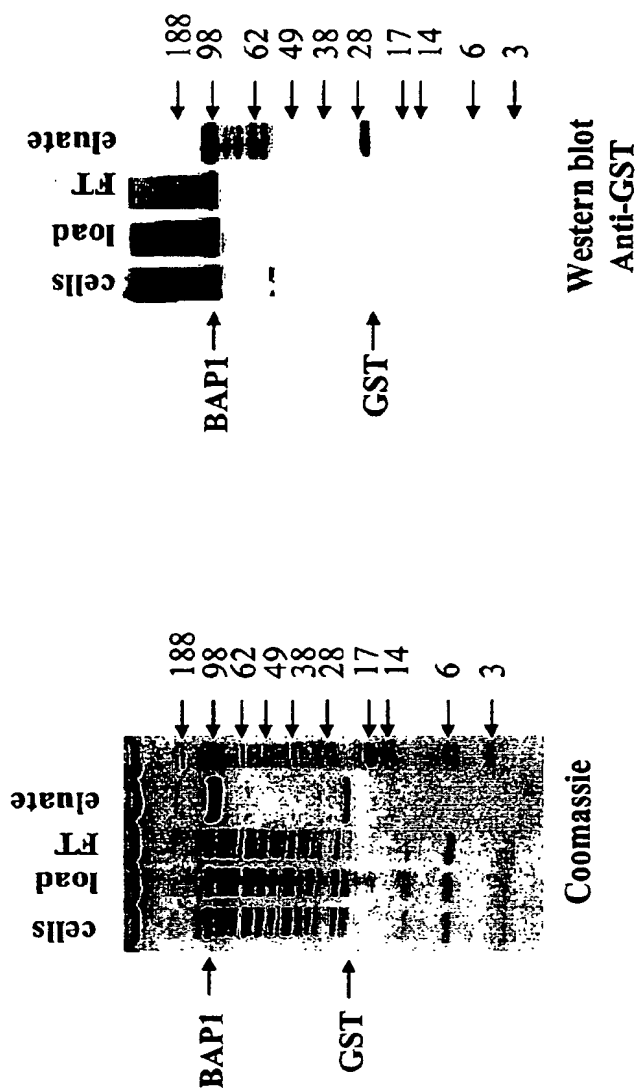
GST-Bap1 was produced using the baculovirus transfer vector pDEST20 along with the Bac-to-Bac baculovirus expression system (Invitrogen). GST-Bap1(1) and GST-Bap1(2) refer to two different virus dilutions used for expression.

FIG. 38

RIGEL

10/510903

BAP1 purification



RIGEL

FIG. 39

Example of Fluorogenic Ub Cleavage Assay

Aminomethyl-coumarin cleavage from Ub C-terminus
generates fluorescence emission in solution-phase assay

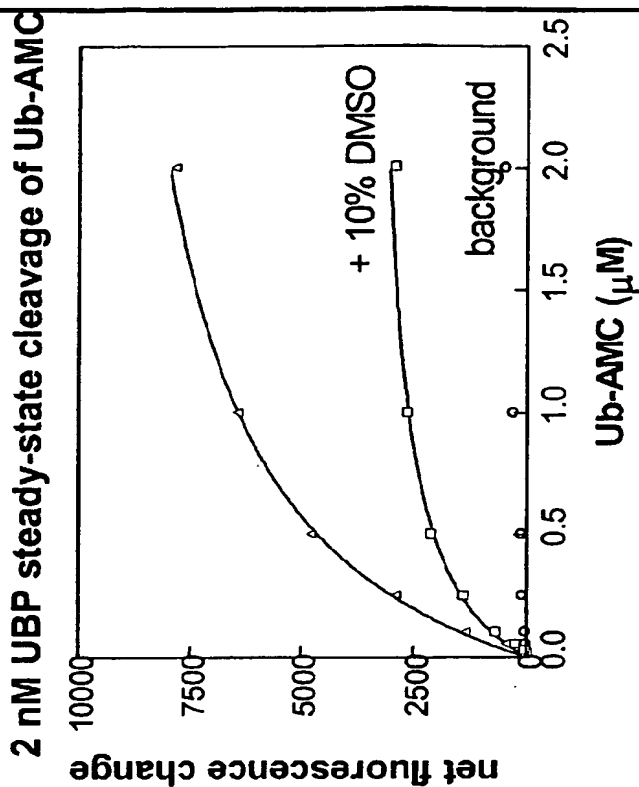
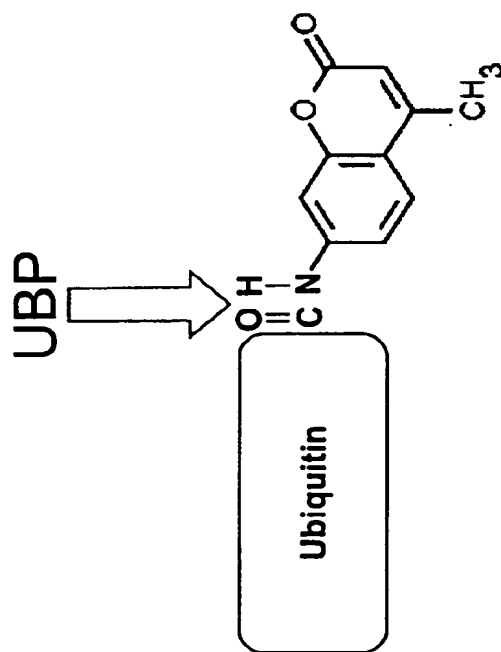


FIG. 40

RIGEL

BAP1 is an Active Ubiquitin Protease

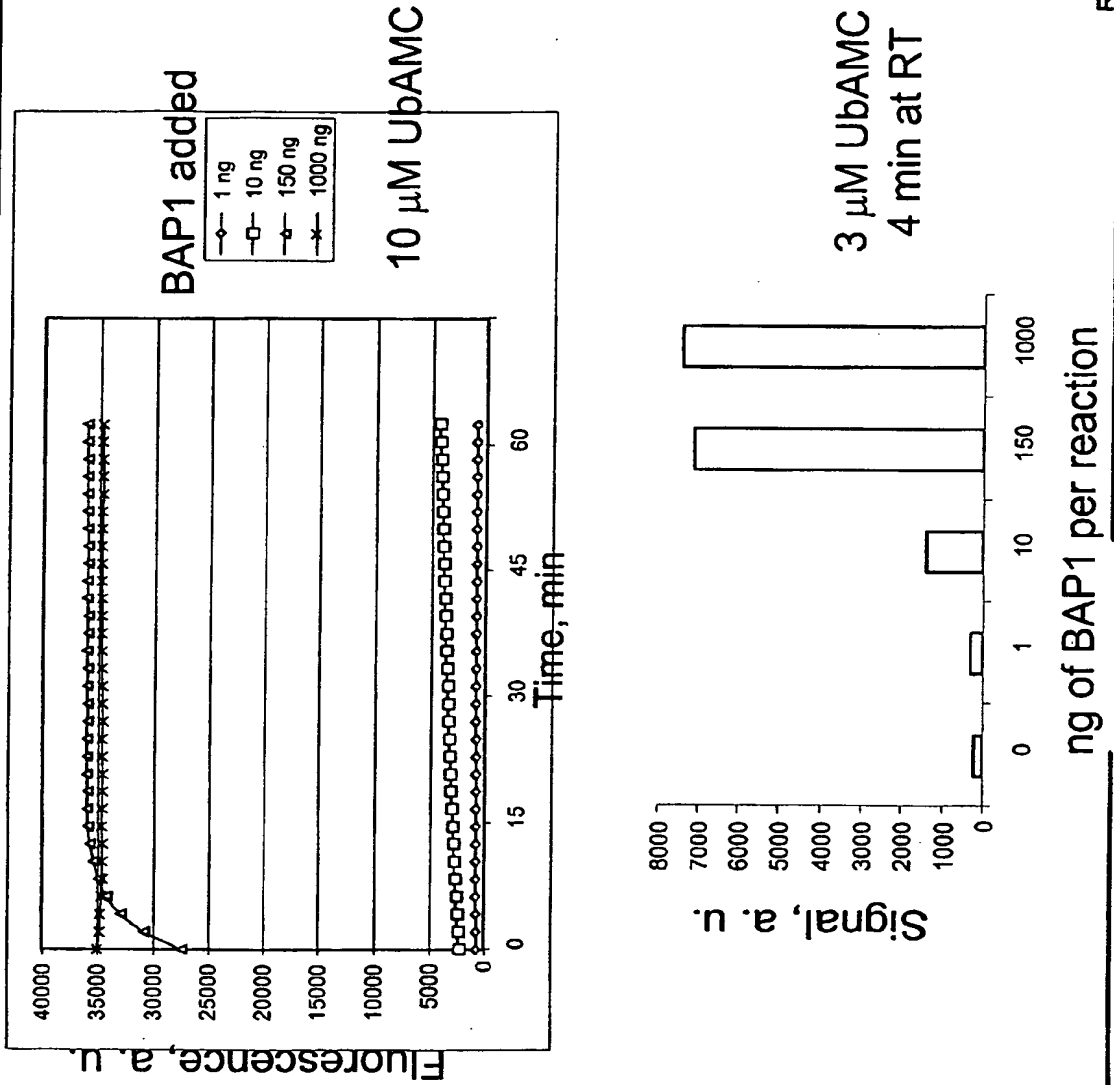
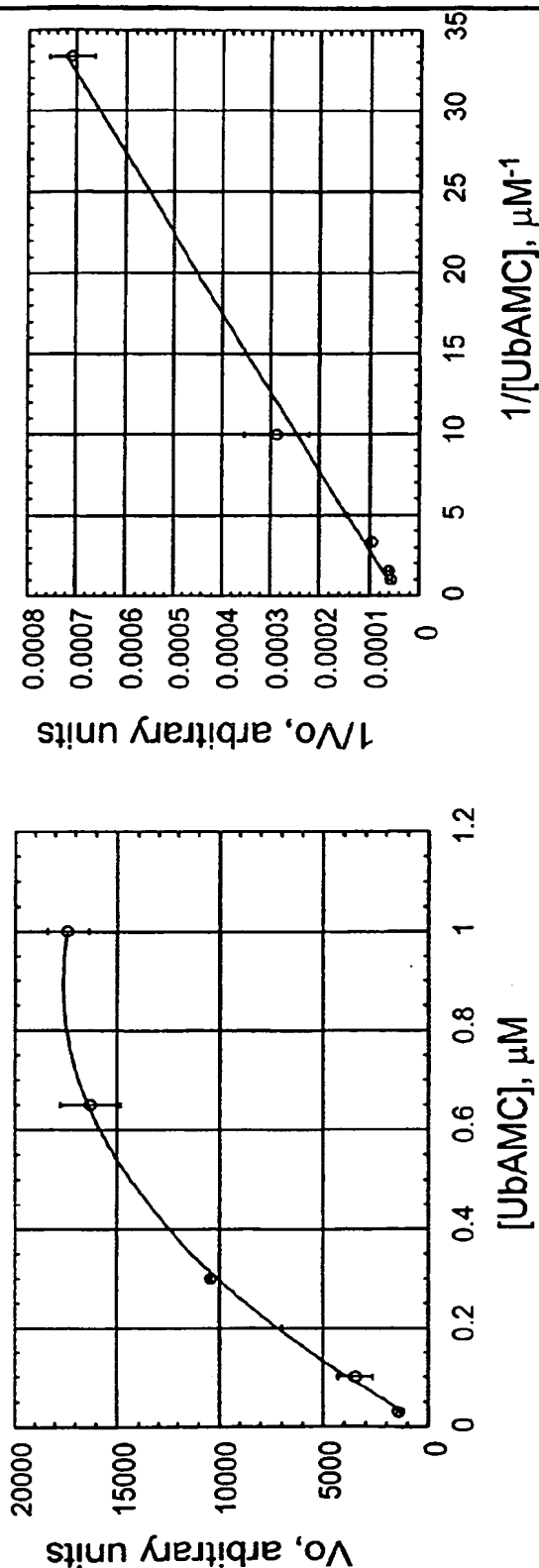


FIG. 41

10/510903

Kinetics of UbAMC cleavage by BAP1



$K_m = 0.5 \mu M$

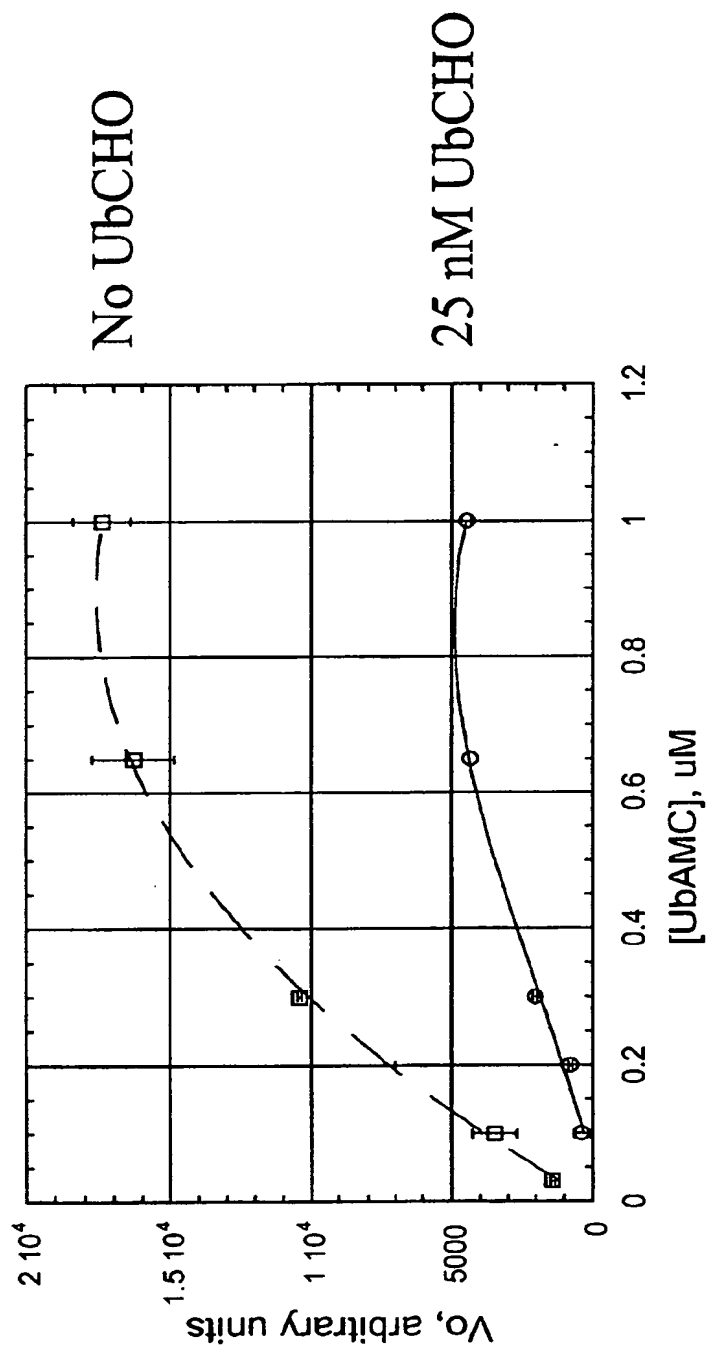
1 ng of BAP1 per well
(0.1 nM BAP1)

FIG. 42

RIGEL

10/510903

UbCHO Acts as Specific Inhibitor of BAP1

 $K_i = 9 \text{ nM}$

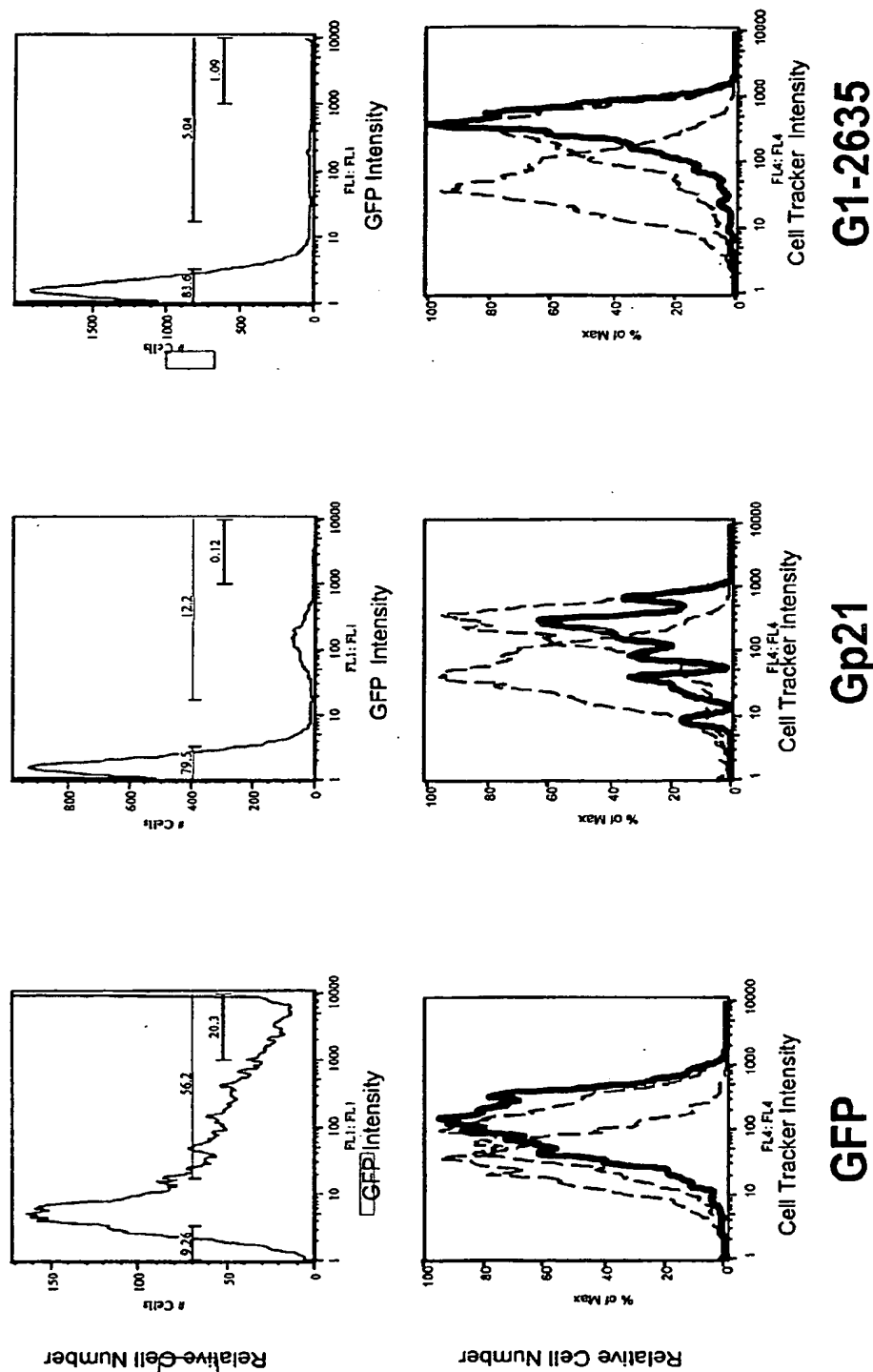
1 ng of BAP1 per well
(0.1 nM BAP1)

FIG. 43

RIGEL

10/510903

The Np95 Functional Hit G1-2635 is Antiproliferative in HMEC Cells



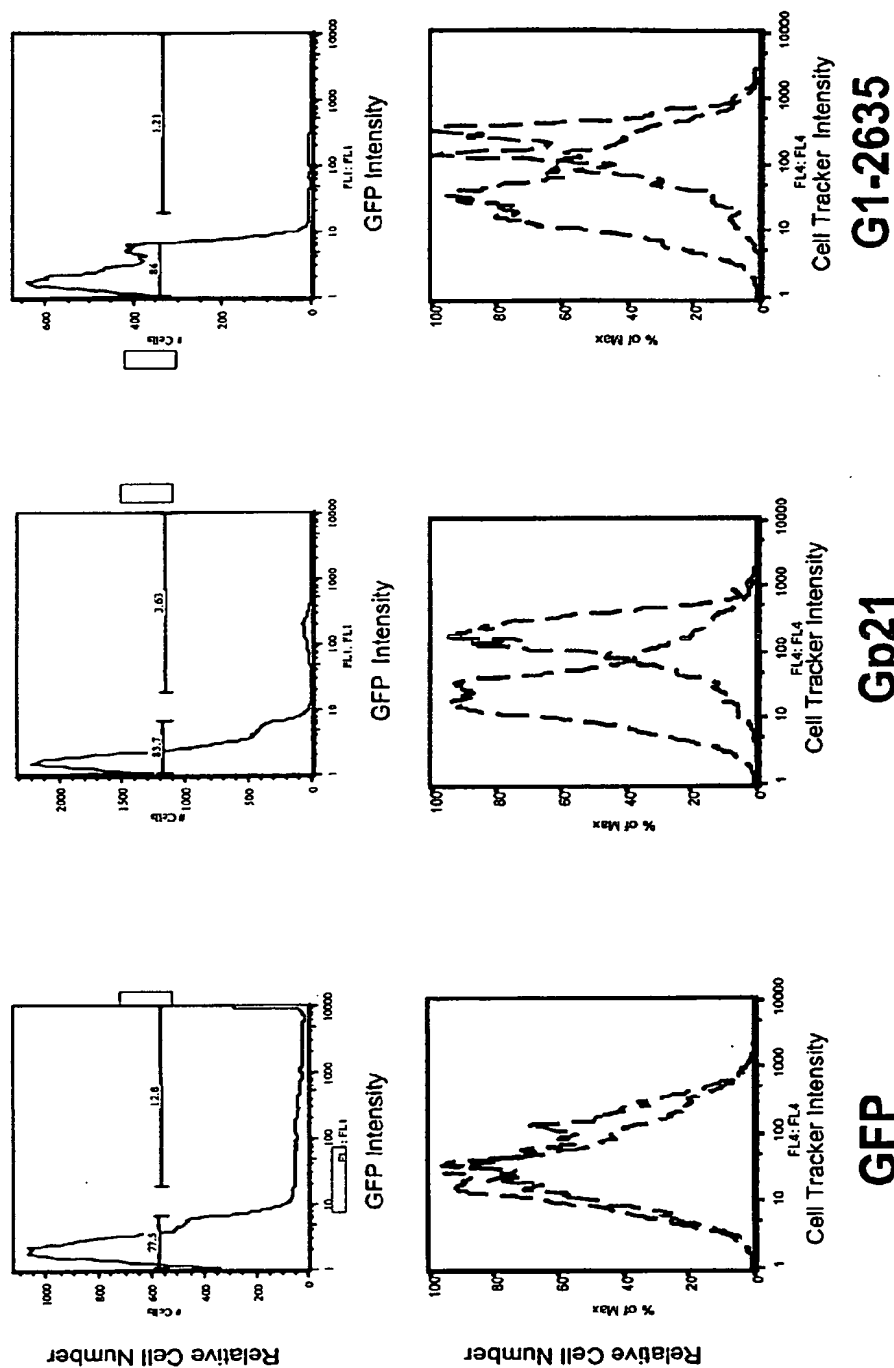
..... GFP+ --- GFP- — GFP HI

FIG. 44

RIGEL

10/510903

The Np95 Functional Hit G1-2635 is Antiproliferative in PrEC Cells



..... GFP+ ---- GFP-

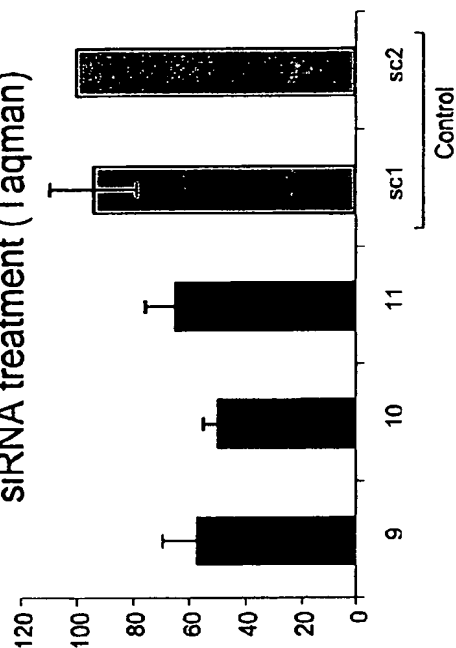
FIG. 45

RIGEL

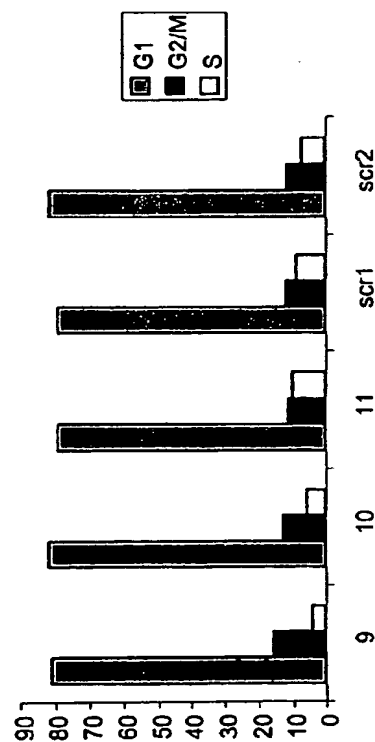
10/510903

NP95 Specific siRNAs Have Antiproliferative Effect on PrECs

NP95 mRNA levels in PrEC after siRNA treatment (Taqman)



PrEC cell cycle profile after NP95 siRNA treatment



BrdU incorporation by PrEC treated with NP95 siRNAs

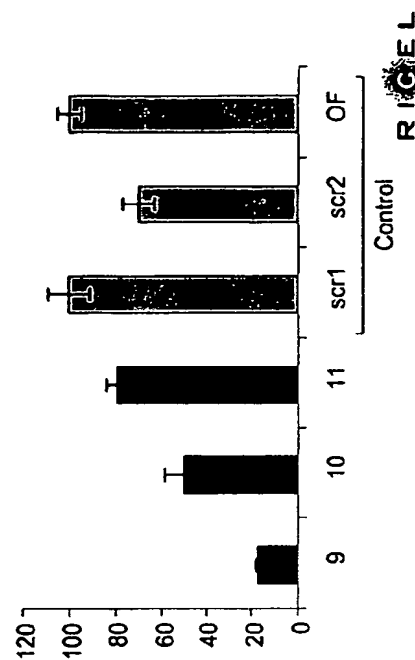
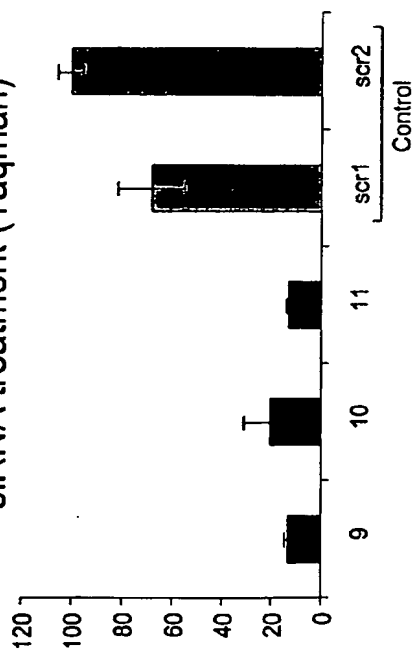


FIG. 46

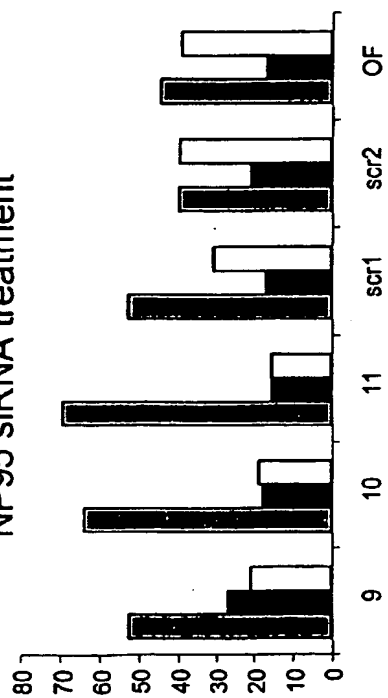
10/510903

NP95 Specific siRNAs Induce G1 Arrest in HUVEC Cells

NP95 mRNA levels in HUVEC after siRNA treatment (Taqman)



HUVEC cell cycle profile after NP95 siRNA treatment



BrdU incorporation by HUVEC treated with NP95 siRNAs

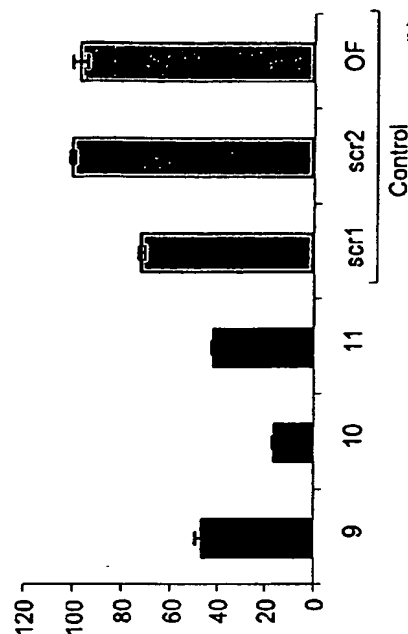
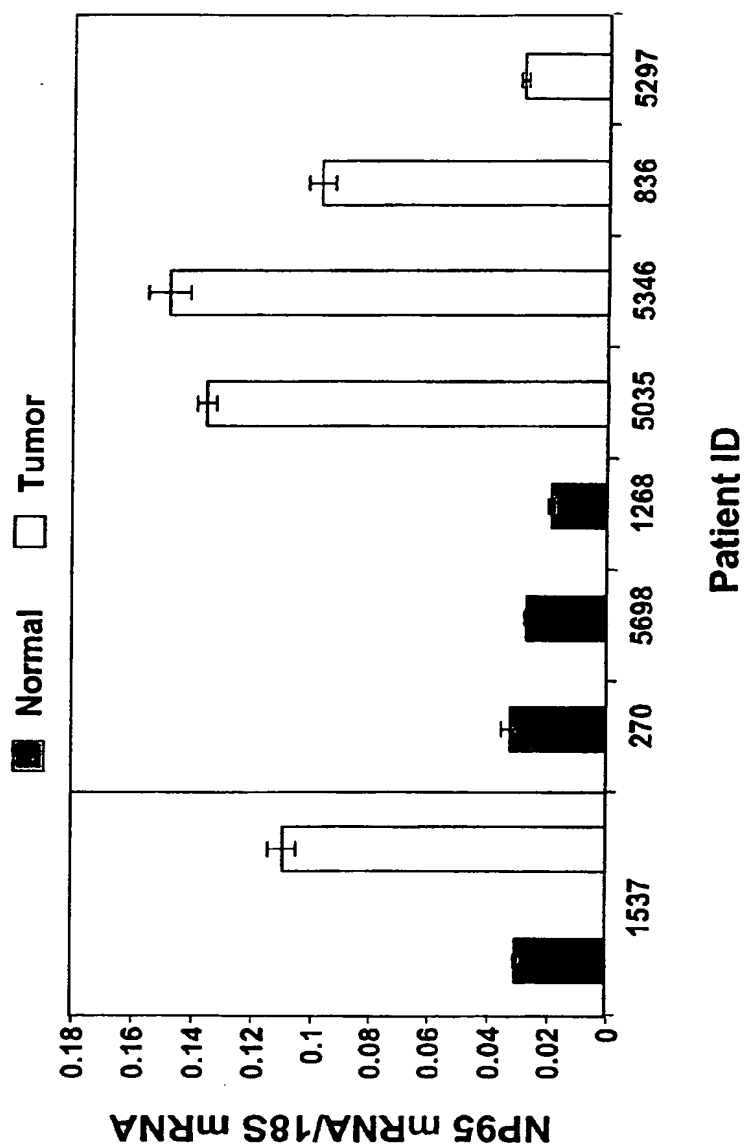


FIG. 47

10/510903

Taqman Analysis of NP95 mRNA Expression in Samples Obtained from Patients with Breast Carcinoma



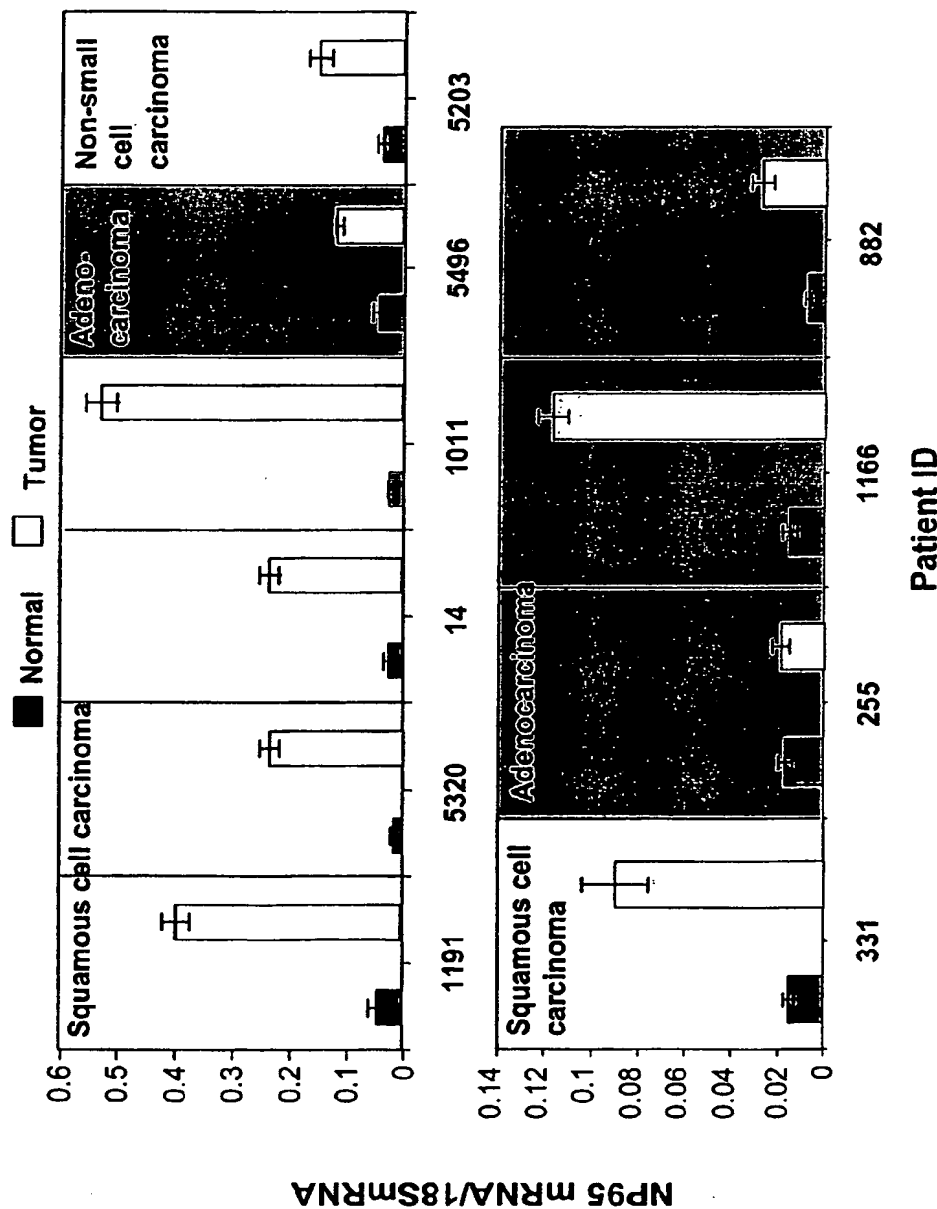
Ductal Adenocarcinoma 836, 1537, 5035, 5346
Lobular Adenocarcinoma 5297

N = 3, 20 ng total RNA/sample

FIG. 48

RIGEL

Quantitative Analysis of NP95 mRNA Expression in Samples Obtained from Patients with Lung Carcinoma



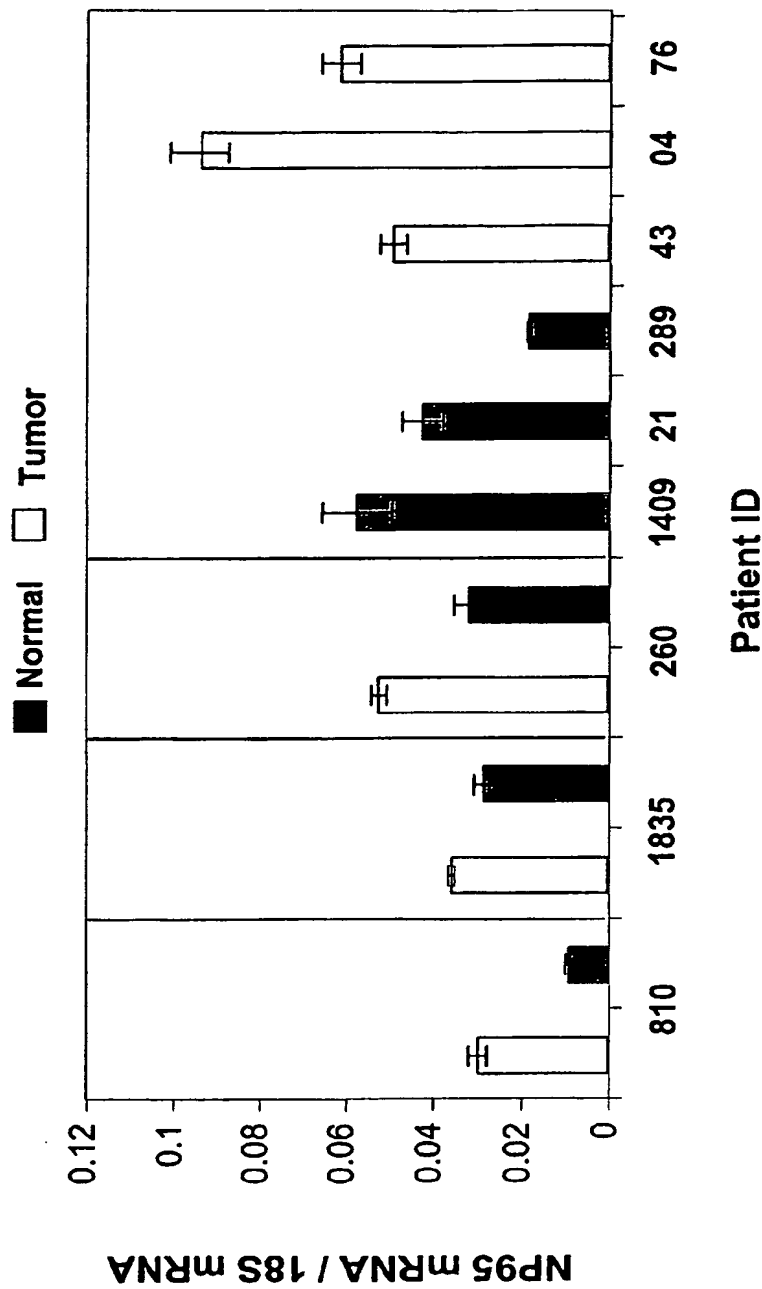
N = 3, 20 ng total RNA/sample

FIG. 49

RIGEL

10/510903

Quantitative Analysis of NP95 mRNA Expression in Samples
Obtained from Patients with Prostate Adenocarcinoma



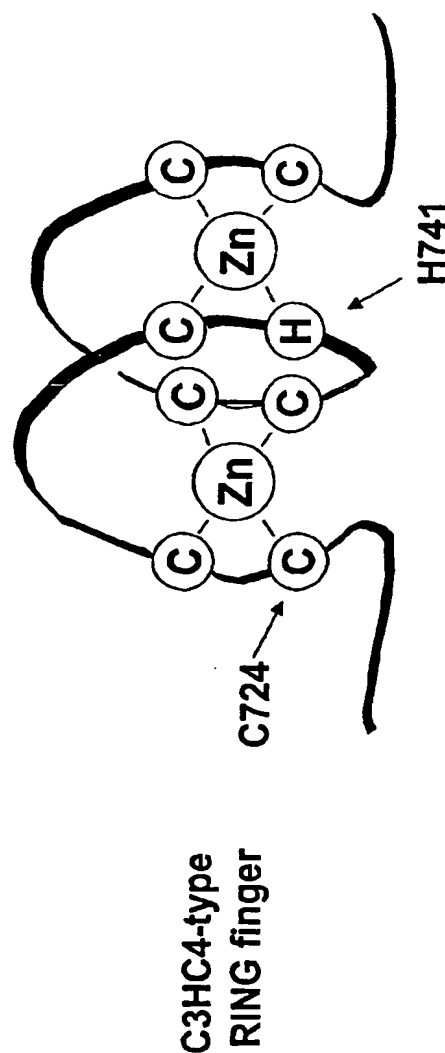
All tumors of acinar cell origin
N=3, 20 ng total RNA Sample

FIG. 50

RI CEL

Dominant Negative Mutants of Np95

| | | | | | |
|------|---------------|-----|-----|-----|------|
| Np95 | WT | UBQ | PHD | G9a | RING |
| | Δ RING | UBQ | PHD | G9a | |
| | C724A | UBQ | PHD | G9a | RING |
| | H741A | UBQ | PHD | G9a | RING |

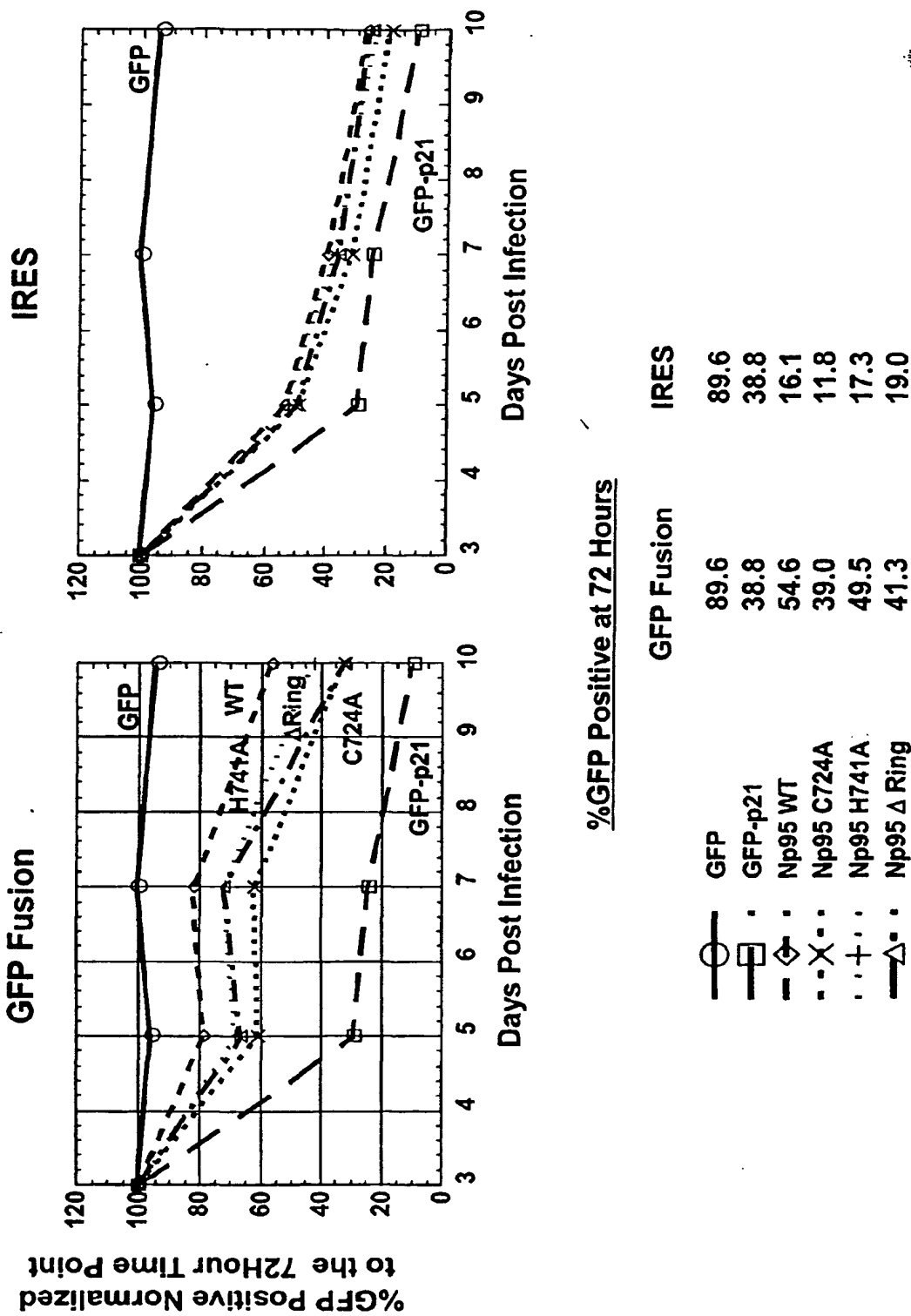


RING

FIG. 51

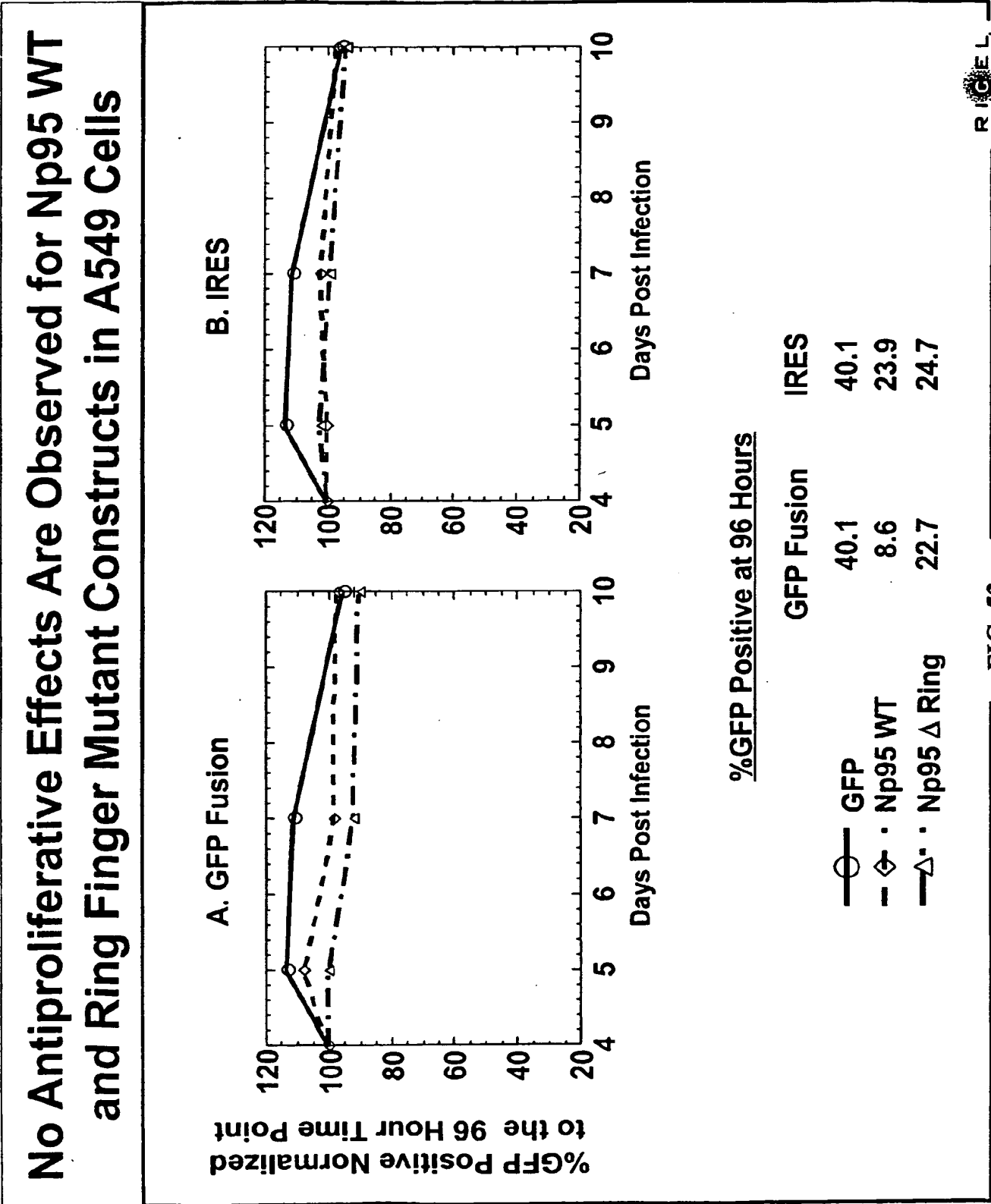
10 / 510903

GFP-fused Np95 Ring Finger Mutants are Slightly more Antiproliferative than GFP-fused Np95 WT in HCT116 Cells

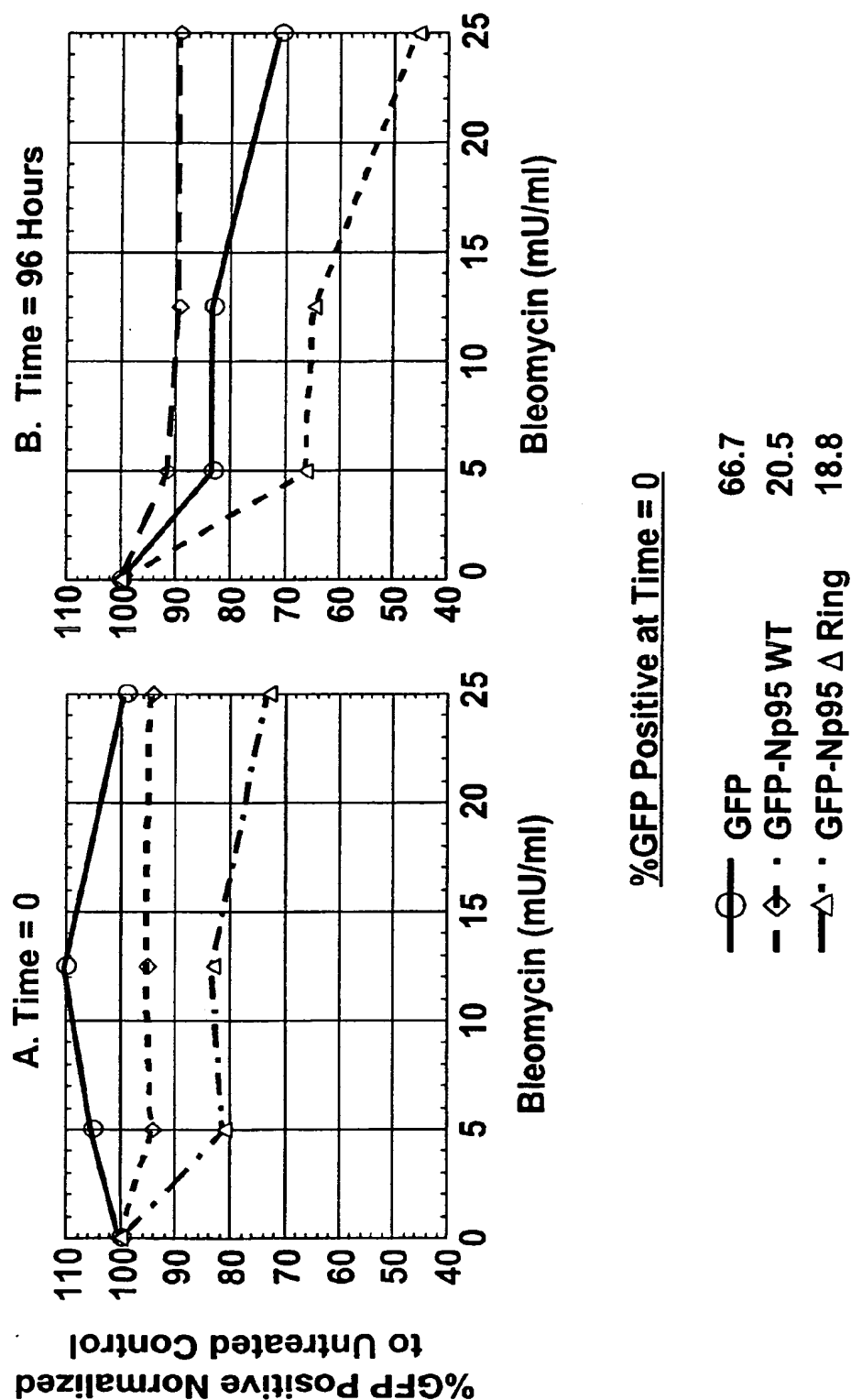


RIGEL

10 / 510903



A549 Cells Expressing GFP-Np95 Δ Ring Become Sensitized to Bleomycin Treatment

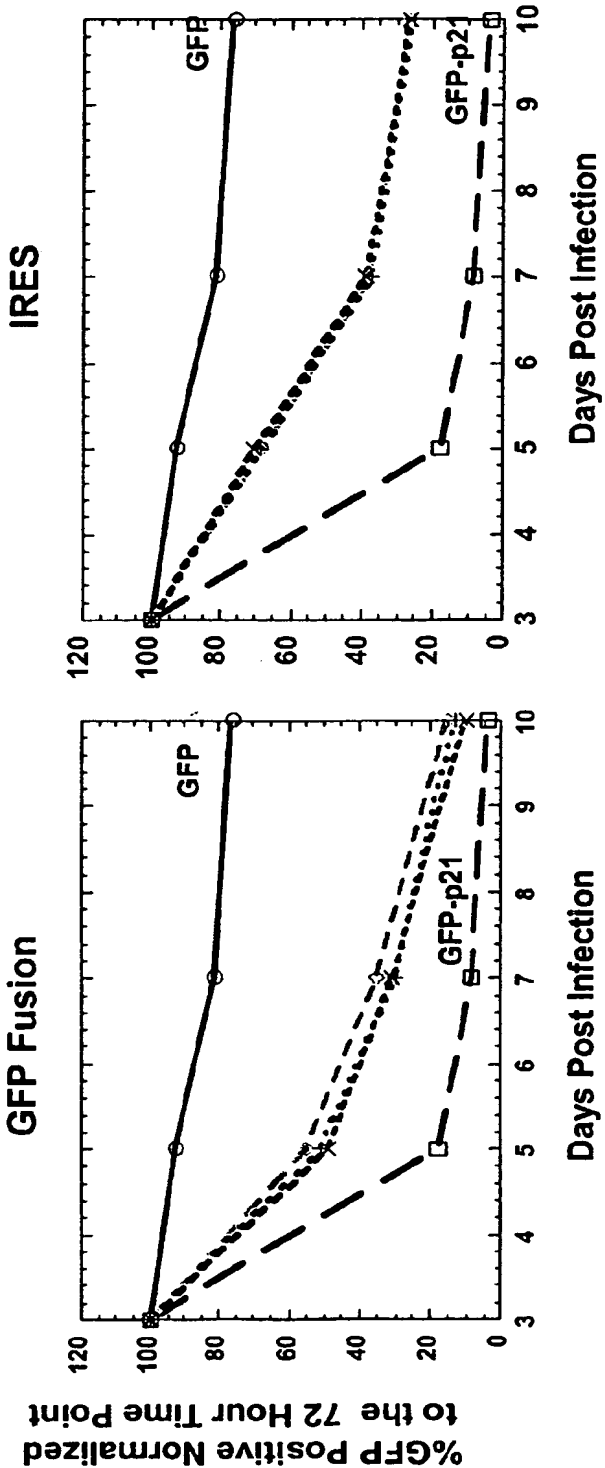


RIGEL

FIG. 54

10/510903

Np95 WT and RING Finger Mutant Constructs are Strongly Antiproliferative in HMECs



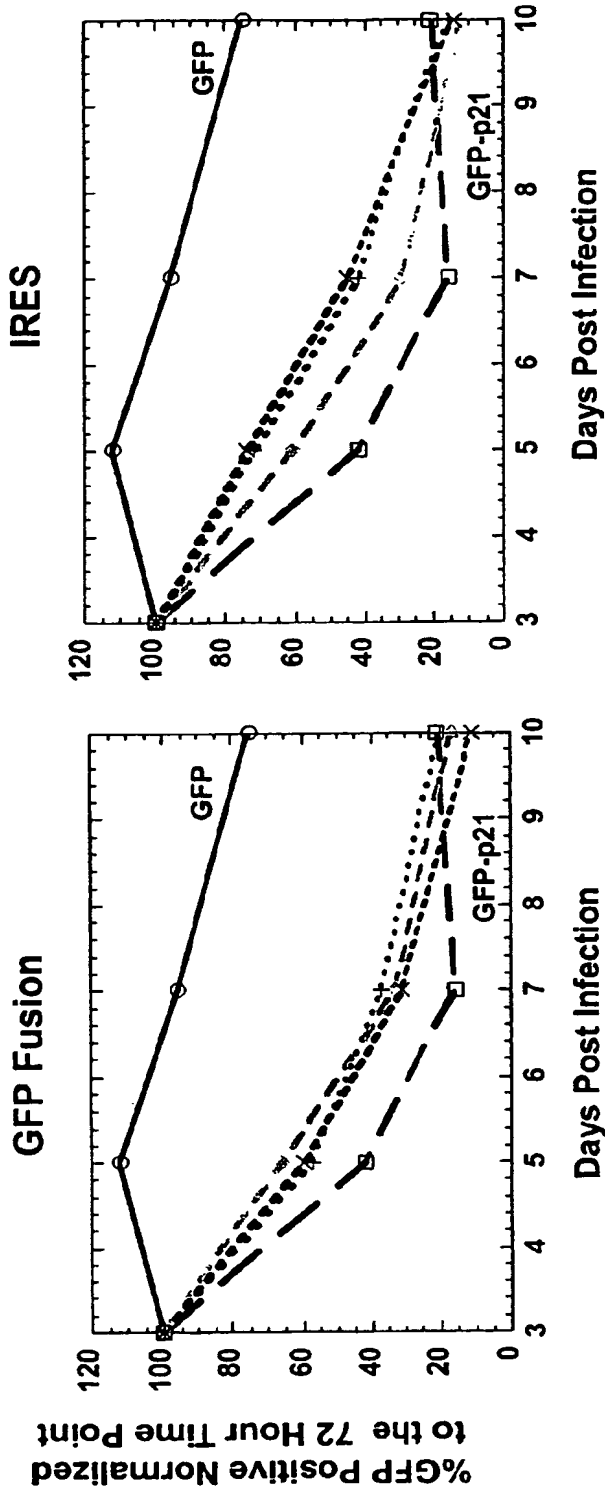
%GFP Positive at 72 Hours

| | GFP Fusion | IRES |
|------------|------------|------|
| GFP | 34.7 | 34.7 |
| GFP-p21 | 7.1 | 7.1 |
| Np95 WT | 5.7 | 10.7 |
| Np95 C724A | 3.5 | 9.8 |
| Np95 H741A | 8.2 | 11.1 |

FIG. 55 RIGEL

10/510903

Np95 WT and RING Finger Mutant Constructs are Strongly Antiproliferative in PrECs



%GFP Positive at 72 Hours

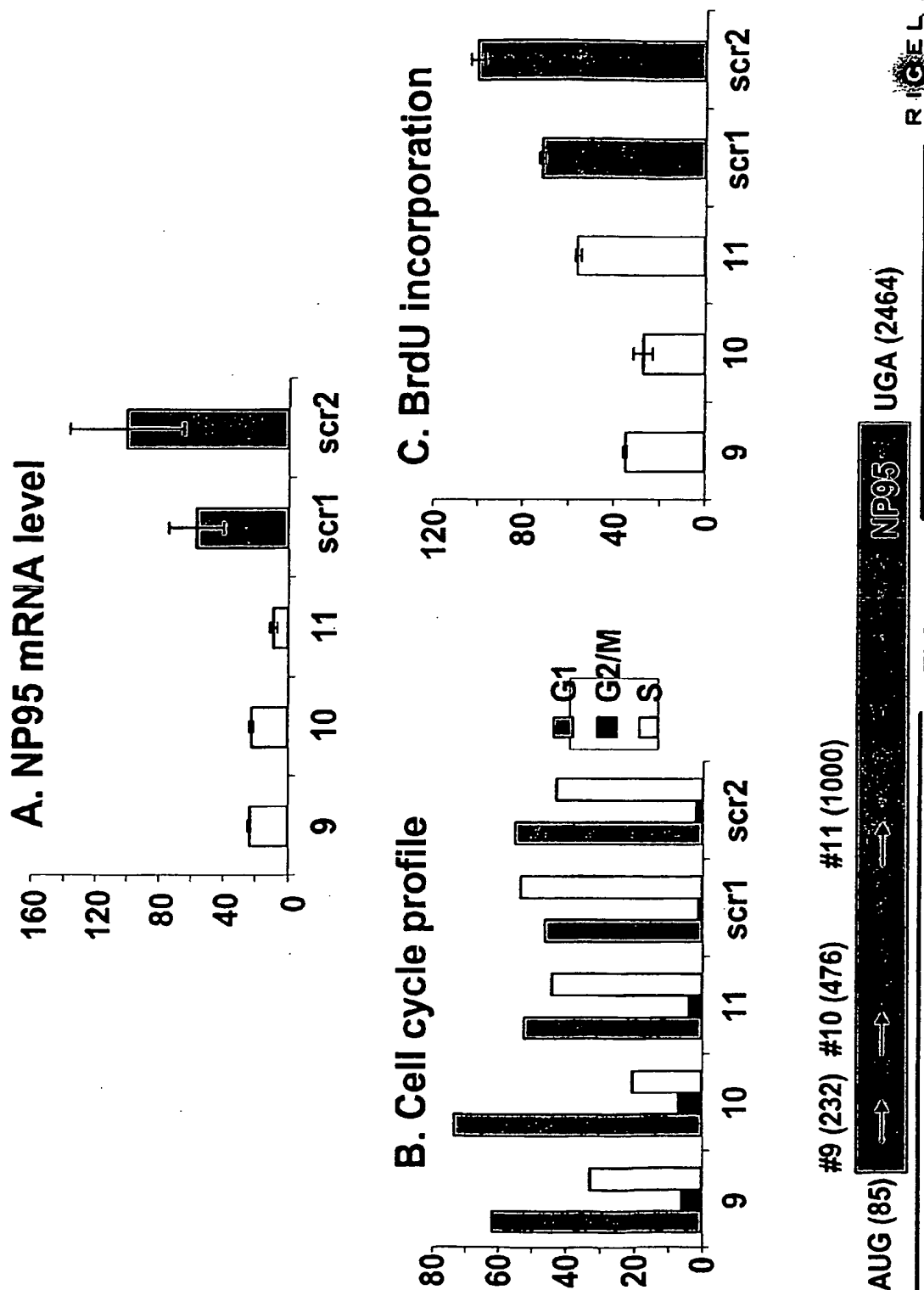
| | GFP Fusion | IRES |
|------------|------------|------|
| GFP | 10.6 | 10.6 |
| GFP-p21 | 1.6 | 1.6 |
| Np95 WT | 1.7 | 2.7 |
| Np95 C724A | 0.91 | 2.3 |
| Np95 H741A | 2.2 | 2.4 |

FIG. 56

RING

10/510903

NP95-specific siRNAs are Antiproliferative in H1299 Cells



The Biochemistry of Ubiquitylation

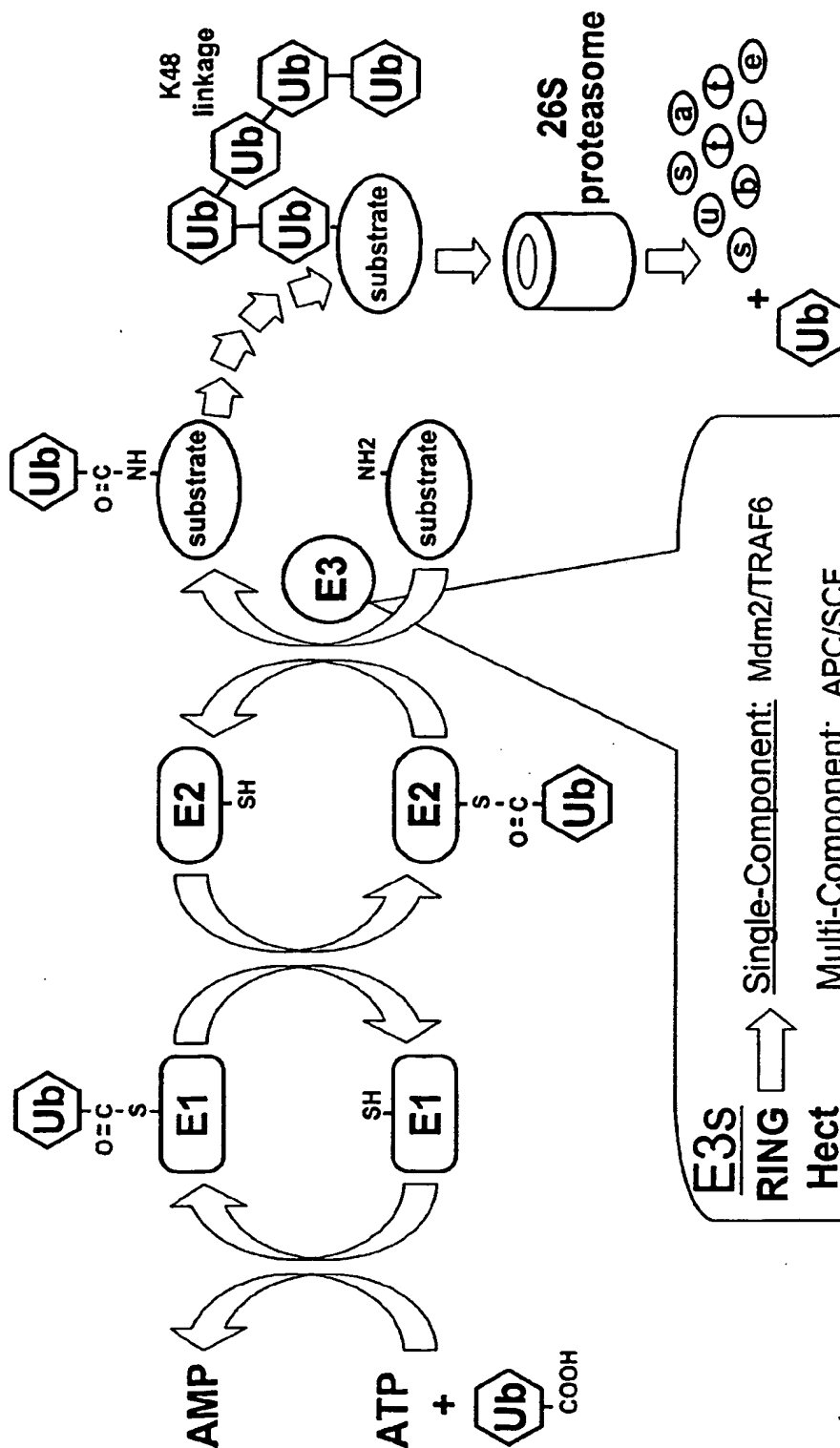
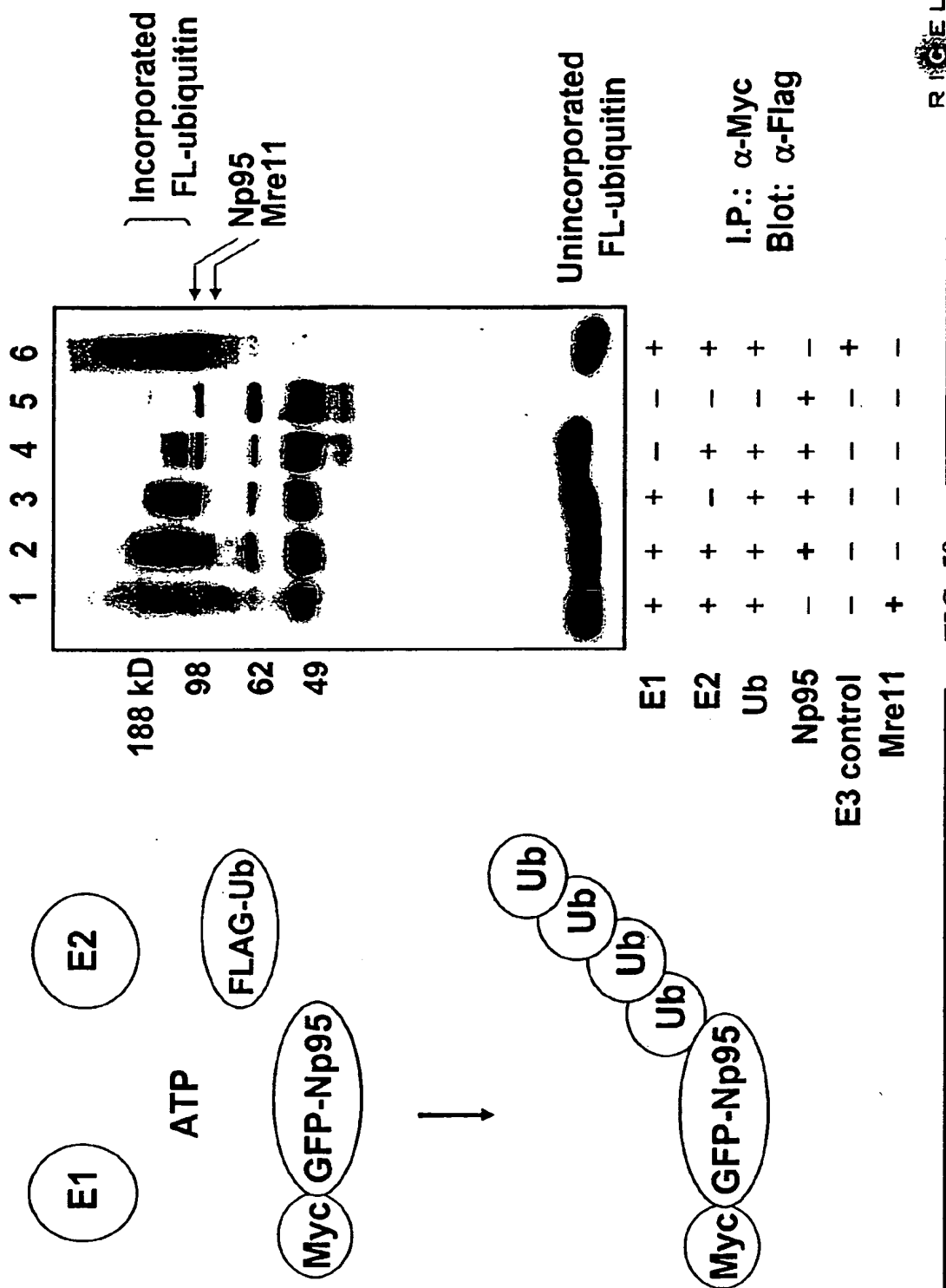


FIG. 58

FIG. 58

GFP-Np95 Exhibits E3 Ubiquitin Ligase Activity



The RING Domain is Required for GFP-Np95 Ligase Activity

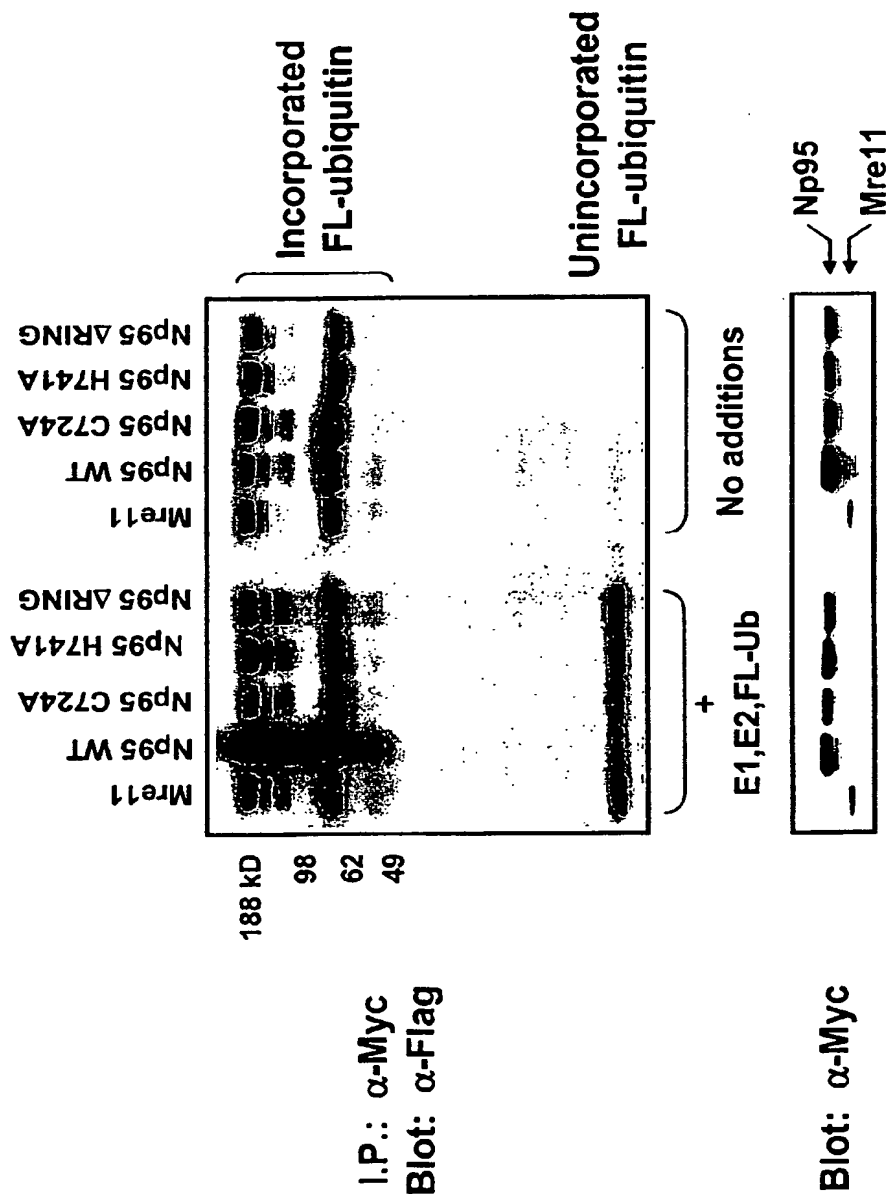
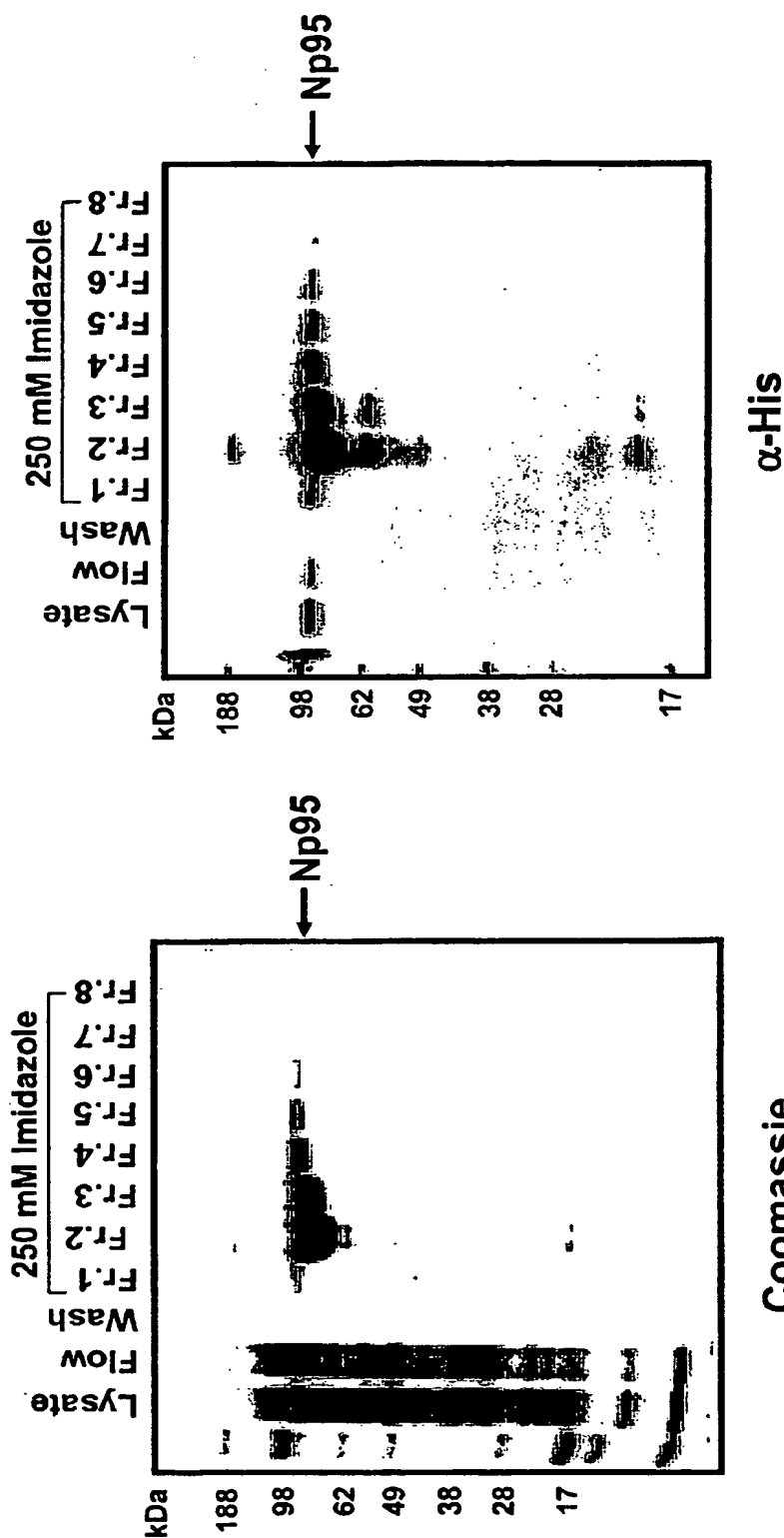


FIG. 60

RIGEL

10/510903

Np95 WT Can be Expressed and Purified from SF9 Cells



Yield: ~2.5 mgs/400 million cells

FIG. 61

RIGEL

10/510903

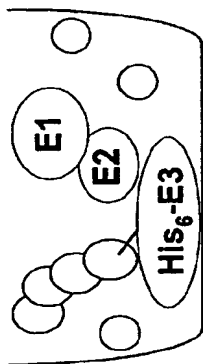
Rigel Plate-Based Ubiquitin Ligase Assay

Nickel coated Plate

Ligase assay reaction
buffer containing Flag-Ubiquitin

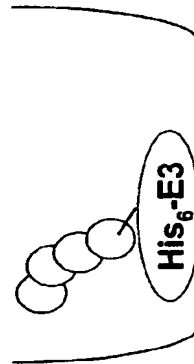
○ Flag-Ubiquitin

↓ Add E1, E2, and His₆-E3



1 Hr. at room temp.

↓ Wash unbound Flag-Ubiquitin



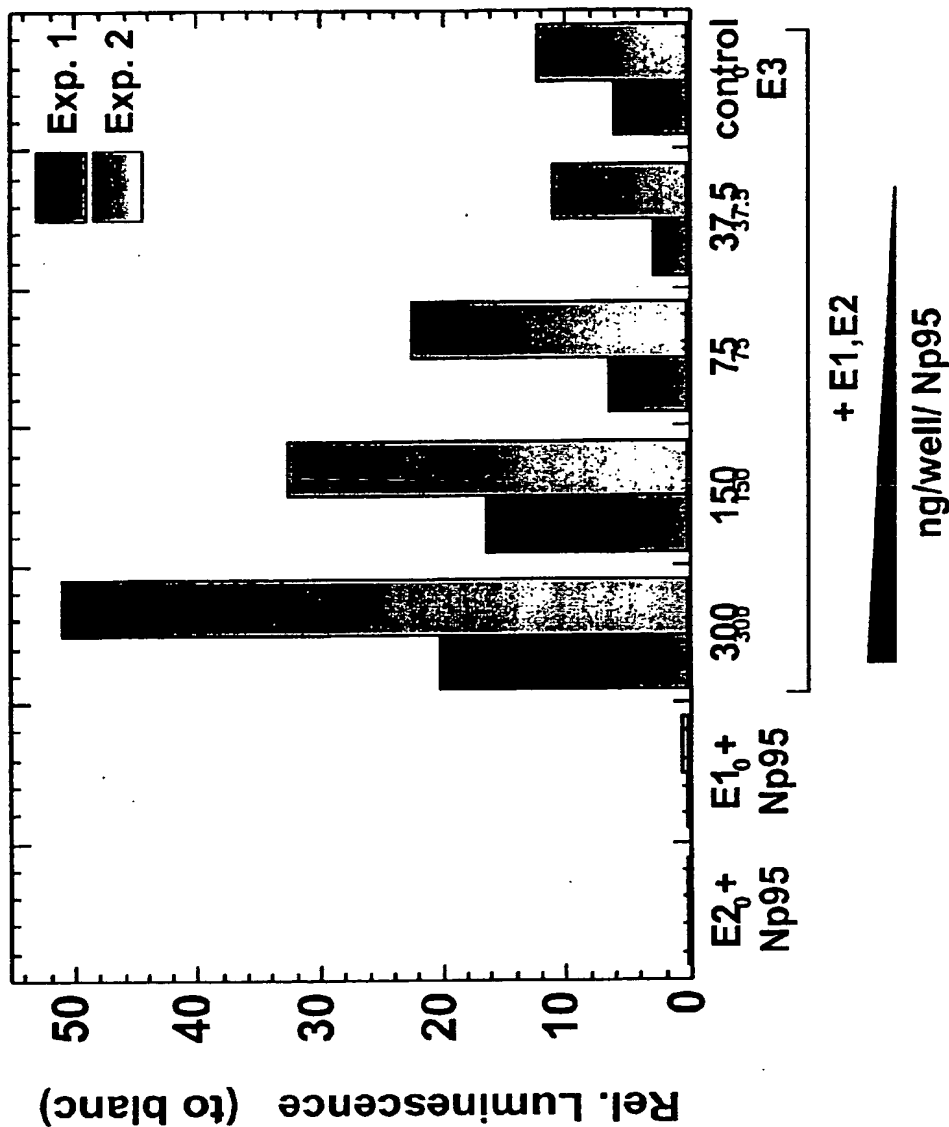
Add α -Flag HRP
conjugate to detect
the Flag-Ubiquitin

Read plate in luminometer

RIGEL

FIG. 62

Np95 Plate-based Auto-ubiquitylation Assay



Reactions contained 100 ng FI-Ub, 5 ng of E1 and, 20 ng of E2 per well. The Np95 controls contained 150 ng Np95. The E3 control contained 75 ng E3. The two data sets are results of duplicate assays.

FIG. 63

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☒ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☒ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.